

# Integrative Survey of Multimodal Analytics, IoT-Enabled Biometric Monitoring, and Artificial Intelligence for Secure, Scalable, and Explainable Healthcare Systems

## Abstract

This survey comprehensively examines the current landscape and future directions of automated, data-driven, and AI-enabled analytics in healthcare, addressing the surging complexity and heterogeneity of clinical, sensor, and IoT-derived data. Motivated by the imperatives of real-time monitoring, precision diagnostics, and personalized therapeutics, the review synthesizes methodological advances across multimodal data fusion, biomedical signal processing, deep learning, and explainable artificial intelligence (XAI). The scope encompasses foundational infrastructures, taxonomy of data modalities, benchmarking datasets, and emergent computational frameworks enabling scalable, privacy-preserving, and interoperable analytics. Key contributions include a critical evaluation of signal processing methods for real-time health monitoring, state-of-the-art multimodal and cross-modal fusion strategies, and adaptive analytics tailored for imperfect, high-velocity IoT streams. The survey elucidates the interplay between synthetic data generation and regulatory, ethical, and technical challenges, emphasizing the centrality of transparency, auditability, and robust privacy/security foundations—particularly within federated and resource-constrained environments. By integrating case studies, comparative paradigms, and application-driven insights, the review highlights best practices in operational analytics, clinical decision support, and the transition to proactive and participatory models of care. Concluding, it identifies persistent challenges—data heterogeneity, annotation scarcity, standardization, equity, interoperability, explainability, and responsible AI governance—charting a coherent roadmap for the transformative, scalable, and ethical deployment of AI-enabled analytics across the biomedical and health sector.

## ACM Reference Format:

. 2025. Integrative Survey of Multimodal Analytics, IoT-Enabled Biometric Monitoring, and Artificial Intelligence for Secure, Scalable, and Explainable Healthcare Systems. In . ACM, New York, NY, USA, 23 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
*Conference'17, Washington, DC, USA*

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM  
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## 1 Introduction

### 1.1 Motivation for Automated, Data-Driven, and AI-Driven Analytics in Healthcare

The escalating complexity of modern healthcare, compounded by the proliferation of vast and heterogeneous data sources, has intensified the imperative for automated, data-driven analytic paradigms. Digital health data now originate from a wide array of modalities, including electronic health records (EHRs), imaging, wearable biosensors, and more. This diversity foregrounds the necessity of capabilities for real-time analysis and continuous monitoring, which are central to advancing personalized diagnostics, adaptive therapeutics, and overall operational efficiency.

Artificial intelligence (AI) and automation facilitate the extraction of actionable insights from high-velocity, complex data streams. These approaches have demonstrated efficacy across domains such as patient safety monitoring, disease surveillance, and optimization of clinical and administrative workflows. Nevertheless, despite these advances, the adoption of synthetic and AI-driven analytics continues to be hindered by several challenges:

- **Practitioner skepticism:** Healthcare professionals remain cautious about entrusting critical decisions to algorithmic systems, citing concerns regarding reliability and interpretability.
- **Regulatory ambiguity:** Established regulatory frameworks have yet to fully accommodate the nuances of automated decision-making in medicine.
- **High stakes of error:** Unlike less critical sectors, errors or malfunctions in healthcare analytics can have profound, direct consequences for patient safety, underscoring the need for robust, transparent, and accountable computational strategies.

### 1.2 Emergence and Significance of Multimodal, Cross-Modal, and IoT-Enabled Platforms

Recent advances in healthcare analytics are increasingly propelled by the integration of multimodal and cross-modal data streams. These encompass imaging, genomics, unstructured clinical narratives, and physiological signals, collectively enabling the construction of comprehensive, patient-specific health models. The deployment of Internet of Things (IoT)-enabled platforms further enhances this landscape by supporting widespread, decentralized data acquisition through interconnected medical devices and in-home sensors.

Such systems are pivotal for continuous, context-aware health monitoring and the facilitation of remote care models. However, their utility is tightly coupled with the ability to harmonize and process data characterized by heterogeneous formats, distinct error

characteristics, and diverse temporal resolutions [29, 68, 78]. Realizing this promise hinges on overcoming several technical challenges:

- **Real-time data fusion:** Integrating diverse data streams in a temporally and contextually consistent manner.
- **Anomaly detection:** Robust identification of clinically significant outliers or data errors across heterogeneous modalities.
- **Adaptive response:** Enabling systems to dynamically adjust analytic or therapeutic strategies in response to incoming data.

Addressing these challenges remains a pressing research frontier that requires not only engineering ingenuity but also methodological innovation.

### 1.3 Central Themes: Big/Synthetic Data, Biomedical Signal Processing, and Real-Time Monitoring

Contemporary innovation in healthcare analytics is driven by paradigms that leverage both big data and synthetic data generation to enhance modeling, simulation, and decision support. Synthetic data, produced using techniques such as generative adversarial networks (GANs), variational autoencoders (VAEs), agent-based simulations, and natural language processing, serves several essential roles:

- **Policy prototyping:** Testing interventions or policies on simulated populations prior to deployment.
- **Privacy risk mitigation:** Reducing re-identification risks when sharing data for research and development.
- **Data augmentation:** Supplementing scarce or biased datasets, thereby improving model robustness and generalizability.

In parallel, advances in real-time biometric monitoring and biomedical signal processing are transforming proactive patient care and population health management. Yet, the reliance on synthetic and AI-generated data introduces new complexities, such as:

- **Definitional ambiguities:** Lack of consensus on terminology and taxonomy for synthetic data categories.
- **Model bias and auditability:** Increased risk of embedding or exacerbating bias, and challenges in tracing model outputs to their data provenance.
- **Privacy challenges:** Uncertainty regarding the adequacy of prevailing frameworks, such as HIPAA or GDPR, to regulate synthetic biomedical data [29].

### 1.4 Scope and Organization of the Survey

This survey synthesizes contemporary research across the domains of multimodal data fusion, real-time biomedical signal processing, and IoT-enabled healthcare systems. Special attention is devoted to the technical, ethical, and regulatory complexities arising from the use of big and synthetic data within biomedical research and clinical practice [29, 68, 78]. The survey critically examines:

- **Methods for scalable analytics:** Including computational frameworks capable of handling high-throughput, heterogeneous healthcare datasets.
- **Privacy-preserving computation:** Covering advances in differential privacy and secure, distributed computation, as well as innovations in digital chain-of-custody mechanisms.

- **Secure data integration architectures:** Focusing on distributed and federated approaches for robust, interoperable data management.

By integrating recent literature and application case studies, this work seeks to elucidate emerging best practices, articulate unresolved challenges, and recommend critical directions for the harmonization of data-driven, patient-centered innovation in the healthcare sector.

## 2 Foundations of Healthcare Analytics and Digital/IoT Infrastructure

### 2.1 Evolution of Digital and IoT Health Systems

The digital transformation of healthcare infrastructure is fundamentally propelled by widespread digitization and the proliferation of Internet of Things (IoT) technologies, which collectively enable the development of highly connected health environments. The adoption of IoT-enabled health systems has facilitated the continuous collection and exchange of real-time patient data via a range of devices, including wearable sensors, mobile health applications, and smart diagnostic equipment. These systems support adaptive, context-aware electronic health services, promoting timely, individualized recommendations that transcend traditional clinical settings [10, 78]. The resulting distributed sensor landscape not only underpins robust health monitoring but also enables the deployment of algorithmically guided interventions, as digital platforms leverage real-time analytics to personalize care and improve the management of chronic diseases [78].

The implementation of digital health technologies—such as health information technology systems, electronic health records (EHRs), and interoperable IoT device frameworks—has significantly expanded both the scope and scalability of healthcare analytics. Advances in federated and cloud-based architectures have facilitated large-scale integration of diverse health data sources, enhancing operational efficiency and supporting advanced clinical research [8, 10, 78, 94]. Contemporary infrastructures enable real-time data ingestion in conjunction with high-throughput analytics. Large healthcare institutions, therefore, can automatically aggregate data streams from clinical, radiological, laboratory, and patient-generated sources, providing the foundation for predictive analytics, surveillance, and personalized medicine initiatives [78, 94]. Nonetheless, the expansion of these systems has exposed persistent inefficiencies, particularly concerning health information exchange, workflow interoperability, and the sustainable deployment of AI-powered applications. These challenges are especially pronounced in rural or under-resourced environments, where resource disparities hinder the effective realization of digital health benefits [10, 94].

A critical examination reveals that the scalability and utility of digital health systems are frequently constrained by non-technical factors. Policy frameworks, for example, exert a major influence on the operational viability and legal boundaries of digital health integration by affecting adoption rates, promoting or hindering interoperability among vendors and platforms, and shaping approaches to patient data governance [8, 10]. Despite technical readiness at the device level, issues related to the lack of standardized protocols, inconsistencies in data representation, and proprietary barriers regularly impede seamless cross-platform data exchange. Ethical

considerations also gain prominence as digitization accelerates, raising urgent issues pertaining to privacy, algorithmic transparency, and the digital divide. The demand for rigorous governance models that prioritize data standardization, promote explainability in algorithmic decision-making, and ensure transparent patient consent is at odds with the current landscape characterized by variable implementation quality and fragmented regulatory environments across jurisdictions [8, 10, 78, 94].

## 2.2 Core Concepts, Data Modalities, and Taxonomies

At the heart of modern healthcare analytics lies the integration of multiple data modalities, which requires a systematic conceptual framework for effective cross-modal data fusion, representation, and interpretation. Key concepts include:

- **Cross-modal learning:** Leveraging relationships between distinct data types for improved inference and knowledge transfer.
- **Multimodal representation:** Encoding and integrating heterogeneous sources—such as imaging, genomics, biosignals, and clinical text—into unified computational models [8, 11, 15, 34, 43, 44, 49, 67, 78, 83, 91, 93, 101].
- **Multi-view fusion:** Aligning complementary perspectives (e.g., EHR text and diagnostic images) to enhance predictive accuracy and model robustness.
- **IoT sensor integration:** Incorporating continuous real-time data capture from sensor streams, thereby promoting context awareness and enabling active feedback mechanisms for both clinical and home settings [8, 49, 67, 78].

Healthcare data modalities are both diverse and complex, necessitating clear taxonomies for rigorous analysis. A comprehensive categorization includes:

- **Omics data:** Genomics, transcriptomics, proteomics, and metabolomics, which characterize molecular profiles.
- **Medical imaging:** Including CT, MRI, PET, and ultrasound, which are foundational for diagnostic and prognostic modeling.
- **Biosignals:** Such as EEG, ECG, and sEMG, offering physiological and functional insights.
- **Clinical data:** Both structured (e.g., lab results) and unstructured (e.g., EHR notes).
- **Behavioral, audio, and video data:** Supporting remote assessment and patient engagement.
- **IoT-derived data streams:** High-frequency, continuous sensor data enabling granular monitoring and timely intervention [8, 27, 28, 35, 47, 51, 53, 61, 69, 77, 78, 87, 102].

The integration of these modalities underscores the increasing analytical complexity of healthcare, with each domain presenting distinct representational, computational, and governance challenges.

Despite significant advances, persistent heterogeneity in data formats, sampling rates, and annotation standards continues to hinder large-scale multimodal analytics. Multi-institutional and multi-vendor datasets are often sparse, fragmented, or marked by

errors and anomalies [8, 10, 24, 78, 94]. Moreover, weak or non-linear correlations between modalities (e.g., loosely coupled sensor streams and clinical events) challenge simplistic fusion strategies and drive the need for sophisticated alignment, imputation, and representation learning methods. Such limitations are exacerbated during real-world deployments, where missing data, class imbalance, and noisy streams are prevalent [24, 78, 94].

Recent developments in cross-modal learning—such as cross-attention mechanisms, graph neural networks with modality-aware encoding, and prompt-based large language models—reflect concerted efforts to close these modality-driven gaps [15, 34, 43, 49, 67, 83, 91]. Nevertheless, key obstacles remain, including the challenge of achieving robust representational alignment under domain shifts, the opacity of current fusion methods, and ongoing concerns about robustness to missing or low-quality modalities. To address these issues, emerging taxonomies and systematic benchmarks of multimodal fusion strategies are proving essential to driving practical advances in healthcare analytics [8, 43, 49].

## 2.3 Datasets, Benchmarks, and Standards

The empirical foundation for digital and IoT health analytics is provided by high-quality, multimodal datasets and robust benchmarking repositories. Several prominent datasets — such as UniMod1K, ImageNet-ESC, LPBA40, IXI, OASIS, ADNI, BraTS, CheXpert, and the MIT-BIH Arrhythmia Database — encompass diverse data types including imaging, biosignals, and longitudinal clinical records, all accompanied by established annotation protocols for algorithmic benchmarking [3, 8, 17, 19, 38, 48, 55, 58, 61, 77–79, 83, 86, 96, 102]. The integration of real-time IoT analytics platforms enables the assessment of algorithms on heterogeneous and high-velocity streaming data, a capability integral to emerging paradigms of remote and continuous care [78, 83].

The range of datasets presented in Table 1 illustrates the breadth of available multimodal resources and their primary roles in benchmarking and advancing algorithmic development.

Achieving harmonization across such datasets remains a significant technical and organizational challenge. Annotation protocols and benchmarking standards must contend with variable data quality and heterogeneous data structures, navigating the tension between rigorous standardization (e.g., radiomics feature definitions and segmentation consistency) and the realities of annotation noise and evolving clinical practices [1, 8, 10, 24, 27, 28, 35, 42, 47, 51–53, 61, 69, 77, 78, 83, 84, 94, 96]. The rise of IoT-specific benchmarking introduces additional performance metrics—such as response ratio, variation factor, and real-time accuracy under fluctuating connectivity or device reliability—that are vital to evaluating real-world analytic performance [78, 94]. Ultimately, the credibility of analytic models is inextricably tied to rigorous data standards, further underscoring the necessity for methodical documentation and consensus-building around feature definitions and provenance.

Despite these advances, notable deficiencies persist. Many datasets remain collected at single sites or within restricted geographic catchments, thus limiting generalizability and frequently neglecting important aspects of population diversity—a particularly consequential shortcoming given the bias sensitivity of contemporary AI-driven models [38, 58, 78]. The rapid diversification of proprietary

**Table 1: Summary of Representative Multimodal Healthcare Datasets and Supported Modalities**

Dataset	Primary Modalities	Data Type(s)	Use Case(s)
UniMod1K	Imaging, Biosignals	Multimodal	Multimodal learning, benchmarking
ImageNet-ESC	Imaging	Images	Image classification, deep learning
LPBA40	MRI	3D Images	Brain segmentation, neuroimaging
IXI	MRI	3D Images	Neuroimaging, brain mapping
OASIS	MRI, Clinical	Images, EHR	Alzheimer's research, diagnosis
ADNI	MRI, PET, Clinical	Images, EHR, Cognitive Scores	Alzheimer's disease progression
BraTS	MRI	Multisequence MRI	Brain tumor segmentation
CheXpert	X-ray	Images, Labels	Chest disease classification
MIT-BIH Arrhythmia	ECG	Biosignals	Arrhythmia detection, biosignal analytics

IoT devices and closed data environments only compounds this fragmentation, introducing inconsistencies in acquisition, storage, and data sharing practices. As a result, emergent multi-center datasets and federated benchmarking initiatives, while promising, require sustainable governance structures, privacy-preserving methods, and equitable resource allocation to realize their full research and translational potential [78, 94].

## 2.4 Data Privacy, Security, and Governance

The exponential scale-up of digital and IoT health systems has elevated the urgency of safeguarding patient privacy, ensuring robust data security, and instituting comprehensive data governance frameworks. EHRs and IoT health platforms are characterized by unique vulnerabilities—such as data breaches, unauthorized access, and failures in transparency—that can undermine patient trust, disproportionately affect marginalized communities, and call into question the ethical legitimacy of digital health programs [10, 11, 24, 34, 35, 40, 44, 45, 47, 51, 52, 60, 61, 73, 78, 79, 89, 90, 93, 94, 98]. Furthermore, systemic inequities in technology access and digital infrastructure exacerbate disparities in the ability to securely access and utilize digital health solutions [78, 94].

National and supranational policy frameworks—including HIPAA in the United States and GDPR in the European Union—define minimum obligations for privacy, data stewardship, and patient rights. Yet, their applicability to fast-evolving IoT-enabled and AI-driven systems continues to be an area of tension and reinterpretation [6, 8, 10, 11, 22, 24, 26–28, 30, 33–35, 38, 40, 41, 44–47, 51–53, 58, 60, 61, 64, 73, 75, 79, 89, 90, 92–95, 98–100, 102]. The distinctive properties of IoT ecosystems—including persistent data generation, dynamic device connectivity, and the involvement of consumer-grade sensors—necessitate specialized privacy-preserving approaches, such as federated learning, on-device analytics, and stringent data minimization, all without compromising analytical utility [8, 24, 94]. Nonetheless, the pace of technological innovation frequently outstrips the evolution of regulatory frameworks, yielding gaps in implementation and leaving ambiguous lines of responsibility among diverse institutional actors.

A comprehensive approach must therefore integrate both technical and procedural innovations in privacy and security with robust governance architectures. Essential elements include:

- Fine-grained, role-based access controls and real-time auditing.

- Formalized consent management mechanisms.
- Transparent algorithmic documentation and reporting.
- Ethical-by-design principles focused on patient autonomy and social accountability.

These requirements are especially challenging for smaller or under-resourced providers, who may lack the technical and organizational capabilities needed for effective compliance [10, 78, 94]. Current best practices and governance models, while taking shape, still fall short of consistently translating regulatory intent into operational reality without hindering innovation or exacerbating inequalities in health access and outcomes.

In sum, the foundational components of digital and IoT health analytics—system infrastructure, data modality integration, standardized benchmarking resources, and privacy-centric governance—must continue to co-evolve. Advancing these domains in a coordinated and ethically robust manner is essential for the realization of equitable, trustworthy, and impactful healthcare analytics.

## 2.5 Biomedical Signal Processing and Real-Time Health Monitoring

**2.5.1 Signal Interpretation and Disease-Specific Applications.** Biomedical signal processing forms the foundation of precise, real-time health monitoring, facilitating sophisticated interpretation of physiological signals such as EEG, sEMG, and ECG. With the advent of advanced quantitative methodologies, disease-specific applications have achieved significant advancements. Notably, intraoperative EEG analysis employing the Hurst exponent as a principal feature has yielded an objective, quantitative means to identify transitions between anaesthetic states, thus enhancing assessment accuracy and providing a critical alternative to subjective clinical judgement during surgery [71]. Such methodological innovations exemplify how the integration of time-series dynamics extends clinical utility.

In the realm of prosthetics, surface electromyogram (sEMG) signals—despite their nonlinear character and susceptibility to noise—have been effectively leveraged for intuitive prosthetic control. Recent developments employ multiresolution decomposition via dual-polynomial interpolation, optimizing denoising and the reconstruction of motor-evoked signals. This approach facilitates reliable multi-class motion decoding within noisy real-world environments, effectively translating complex biosignals into responsive, user-adaptive prosthetic commands [80]. The robustness of

these preprocessing pipelines is central to the real-time applicability of prosthetic systems.

The computational analysis of ECG signals further illustrates the transformative impact of advanced processing strategies on disease detection. Transforming one-dimensional ECG traces into two-dimensional time-frequency representations—particularly using the continuous wavelet transform (CWT)—has measurably improved arrhythmia classification accuracy. Through the application of transfer learning and pre-trained convolutional neural network architectures, such as ResNet-18, classification accuracy rates up to 96.17% have been achieved on benchmark datasets, especially when employing Ricker wavelet-based transforms. These state-of-the-art algorithms not only deliver high efficiency but also offer interpretability via visualization techniques such as Grad-CAM, thereby facilitating clinician engagement and review [86].

Reliable benchmarking is underpinned by widely recognized datasets, including MIT-BIH for ECG, NinaPro for sEMG-based prosthetic studies, and CHB-MIT/other EEG repositories for neurological disorder monitoring [71, 80, 86]. These curated datasets serve as reference standards, bolstering reproducibility and standardization essential for rigorous comparative algorithmic development and translational research.

Despite these advances, significant challenges persist in the real-time interpretation of multimodal signals, particularly in practical settings involving data streams from wearable and IoT-enabled platforms. The integration of heterogeneous data sources, each with varying sample rates, modalities, and transmission reliability, complicates timely and accurate monitoring—especially during acute clinical scenarios [2, 37, 49, 78]. Clinical applications such as seizure and home-based epilepsy monitoring are emblematic of these challenges, introducing artifacts, motion-induced noise, and missing data that necessitate advanced artifact rejection and robust adaptation strategies.

Innovative multimodal systems exemplify the cutting edge of real-time monitoring. Advanced seizure detection platforms combine multiresolution analysis (e.g., maximal overlap discrete wavelet transforms, MODWT) and deep learning, employing patient-specific models to autonomously decompose EEG signals, extract hierarchical features, and aggregate predictions for sensitive, low-false-alarm alerts—functioning effectively even in ambulatory, uncontrolled environments [2, 18, 37]. The integration of wearable EEG with inertial measurement units (IMUs) further enhances detection reliability by incorporating movement context, enabling reliable differentiation between neurophysiological events and artifacts. The synergy between multimodal feature extraction and intelligent artifact rejection supports continuous monitoring beyond traditional clinical settings, representing a pivotal stride toward pervasive digital health.

**2.5.2 Advanced Feature Learning and Biometric Monitoring Algorithms.** The progression of feature learning paradigms has markedly improved both the accuracy and interpretability of biometric monitoring systems. Deep and hybrid architectures—including recurrent neural networks (RNNs) and long short-term memory networks (LSTMs)—successfully capture temporal dependencies inherent in biomedical time series. Building on these, frameworks such as the Hybrid Recurrent Long Short-term based Tyrannosaurus Search

(HRLS-TS) have emerged, providing adaptive, real-time monitoring and superior signal discrimination in dynamic conditions [49]. The deployment of bio-inspired metaheuristics, as in the Tyrannosaurus Search optimizer, further enhances algorithmic performance by balancing predictive accuracy with computational constraints.

A noteworthy evolution within the field is the shift toward interpretable, multi-view feature integration. The Advanced Multi-View Deep Feature Learning (AMV-DFL) framework epitomizes this trend, amalgamating conventional frequency- and time-domain features with automatically-learned deep representations. This approach delivers the performance advantages of deep learning while retaining the transparency crucial for clinical acceptance. Methods such as SHAP (Shapley Additive Explanations) facilitate rigorous model interpretability by quantifying each feature's or channel's contribution, supporting clinical decision processes and addressing regulatory requirements [2, 49]. The imperative for explainable models persists as real-time monitoring solutions transition from laboratory implementation to patient-facing clinical tools.

Real-time signal optimization strategies remain indispensable, particularly as wearable and IoT platforms are often constrained by limited power and processing resources. These adaptive strategies balance model complexity, energy efficiency, and computational latency, thereby safeguarding the reliability and integrity required for continuous, personalized monitoring in modern healthcare scenarios [49].

**2.5.3 Evaluation and Real-World Validation.** The successful translation of biomedical signal processing research into large-scale, real-world health monitoring depends on robust evaluation protocols. As IoT-enabled healthcare grows, reliance on traditional accuracy metrics alone has become insufficient. To address operational realities, a range of performance indicators—including service response ratio, delivery time, variation factor, identification ratio, and aggregate processing time—are regularly employed to assess system viability under diverse streaming data and network conditions. These metrics permit comprehensive appraisal of a platform's responsiveness and resilience in the face of fluctuating loads, real-world device failures, and transient data disruptions [78].

The adoption of such metrics (Table 2) allows for critical insight into the operational strengths and limitations of signal monitoring frameworks.

A central priority in contemporary validation efforts is the utilization of diverse, multicenter, and authentic clinical datasets. Data amalgamated from large-scale hospital infrastructures and wearable device deployments ensures sufficient heterogeneity for comprehensive benchmarking, exposing analytic algorithms to an array of physiological, demographic, and environmental confounders [8, 18, 26, 49, 54, 77, 94]. This strategy strengthens model generalizability and facilitates identification of persistent challenges, including missing data, context-dependent adaptation, and site-specific bias—issues increasingly addressed through federated and transfer learning approaches.

Artifact rejection and adaptive recalibration remain vital to reliable deployment. Notably, multicenter evaluations in seizure detection demonstrate that the combination of auxiliary sensing modalities (e.g., IMUs) and advanced postprocessing significantly reduces false alarm frequency without compromising sensitivity [18]. The

**Table 2: Representative Metrics for Evaluating Real-Time Health Monitoring Systems**

Metric	Description
Service Response Ratio	Measures the fraction of health events correctly recognized and responded to within a defined period
Delivery Time	Quantifies time taken from signal acquisition to response generation
Variation Factor	Assesses system performance stability under changing network or signal conditions
Identification Ratio	Evaluates correct identification of target events relative to all system outputs
Aggregate Processing Time	Captures total computational time for processing and classification tasks

growing integration of explainable AI techniques further enhances clinician trust, providing actionable insights and facilitating safe and effective intervention.

In summary, the discipline is transitioning from incremental algorithmic improvements to the deployment of comprehensive, integrated pipelines that emphasize signal fidelity, interpretability, computational efficiency, and operational robustness—collectively supporting the broader adoption and real-world impact of real-time health monitoring technologies.

### 3 Cross-Modal, Multimodal, and IoT-Driven Healthcare Analytics

#### 3.1 Proportionate Data Analytics (PDA) and Data Management

The rapid expansion of Internet of Things (IoT) devices within healthcare ecosystems has culminated in the generation of voluminous, heterogeneous, and high-velocity data streams. These streams frequently exhibit variable formats, fluctuating quality, and diverse error profiles—characteristics that challenge traditional analytic frameworks, which implicitly assume uniform reliability among data sources. Such conventional models prove inadequate in environments where data quality is nonstationary and prone to both transient anomalies and persistent sensor failures.

Proportionate Data Analytics (PDA) offers a principled paradigm for addressing these complexities. By employing statistical techniques—for example, linear regression applied to temporally disjoint intervals—PDA frameworks can discriminate between routine signal variability and substantive anomalies, thus enabling the rapid identification and isolation of compromised data streams without excessively impacting system responsiveness or service continuity [78]. This context-sensitive approach allows healthcare systems to dynamically tailor data processing pipelines: streams exhibiting anomalous behavior, incomplete information, or degraded quality are proportionately de-emphasized or routed for secondary verification, while reliable signals maintain their operational priority. In practice, this strategy enhances overall analytic robustness and ensures that clinical decision-support systems remain well-calibrated, even in the face of environmental noise and transient uncertainty.

The imperative for PDA in healthcare stems from two principal motivations:

- **Mitigating Downstream Analytic Risk:** To ensure that decision-support mechanisms are governed by the most reliable data, thereby minimizing the possibility of error amplification and unintended clinical consequences.

- **Autonomous Adaptive Resource Allocation:** To empower monitoring platforms and autonomous diagnostic systems with the capacity to allocate analytic resources dynamically, guided by real-time assessments of data quality and provenance.

Crucially, PDA frameworks embed anomaly detection, quality scoring, and error flagging directly into the data ingestion and management layers, thereby maximizing operational transparency. By systematically documenting points of error, correction, and exclusion throughout the analytic workflow, PDA not only serves technical robustness but also aligns with regulatory mandates for traceable data lineage and auditability in clinical contexts.

#### 3.2 Multimodal and Multisource Data Fusion

The confluence and synthesis of diverse biosignals, medical images, behavioral metrics, and IoT-derived contextual data have revolutionized the scope and depth of healthcare analytics. Integrative multimodal and multisource fusion methodologies are indispensable for leveraging the distinctive yet complementary informational content inherent to each modality, whether in continuous physiological monitoring, static or temporal imaging, longitudinal behavioral profiling, or the incorporation of structured electronic health records [8, 27, 28, 30, 31, 33–35, 42, 47, 49, 51, 53, 56, 61, 77, 78, 87, 94]. Among canonical applications are:

- **Radiomics:** Extraction and fusion of high-dimensional features from CT, MRI, or PET imaging with omic datasets and clinical histories, underpinning enhanced biomarker discovery and risk stratification.
- **Cardiometabolic Monitoring:** Integrated use of wearable biosensors with behavioral and environmental data for robust, longitudinal risk assessment.
- **Intelligent Hospital Platforms:** Cohesive synthesis of laboratory results, procedural logs, and real-time monitoring to support advanced analytics, AI-driven triage, and operational optimization.

To effectively harness multimodal data, several distinctive fusion paradigms have emerged:

Systems frequently require dynamic selection among these paradigms, dictated by the evolving data landscape, operational constraints, and analytic objectives (see Table 3).

A central technical challenge in this context remains the quantification and propagation of uncertainty, which is vital for trustworthy diagnostic inference and risk management in clinical settings. Approaches such as Dempster-Shafer theory have been adapted

**Table 3: Comparison of Multimodal Data Fusion Paradigms**

Paradigm	Characteristics	Advantages and Limitations
Early Fusion	Merges raw features from all modalities before modeling; enables low-level cross-modal interactions	Captures direct synergies; prone to dimensionality issues, may require large datasets
Late Fusion	Combines outputs from modality-specific models (decision-level)	Robust to missing modalities; may overlook deep synergistic structure
Joint/Hybrid Fusion	Interleaves intermediate representations, often with attention mechanisms or layered interactions	Balances representation richness with tractable modeling; excels in streaming and heterogeneous IoT environments

to healthcare data fusion to aggregate evidential support, capturing both the strength of belief and the degree of conflict among modalities [25]. These probabilistic frameworks facilitate a nuanced response to incomplete or contradictory data. Additional strategies for addressing missingness and partial observation include:

- **Domain-Informed Statistical Imputation:** Leveraging clinical knowledge for plausible data reconstruction.
- **Representation Learning with Variable-Length Sequences:** Designing models that are robust to interrupted or incomplete input streams [6, 8, 26, 77, 94].

Collectively, these innovations constitute the methodological backbone of highly resilient, data-rich analytics addressing real-world healthcare complexity.

### 3.3 Emerging Adaptive and Cross-Modal Processing

The dynamic, often multi-institutional, and privacy-sensitive character of contemporary healthcare data ecosystems has driven the advancement of learning paradigms that are adaptive, distributed, and effective under limited annotation. In distributed IoT healthcare, adaptive, semi-supervised, and federated multi-source fusion strategies have demonstrated substantial promise: enabling analytics at scale even when labeled data is scarce, privacy regulations are restrictive, and data distributions drift over time [49, 63, 78].

- **Self-Supervised Learning:** Facilitates the extraction of generalized data representations from large unlabeled corpora, allowing efficient transfer to downstream clinical tasks.
- **Continual Learning:** Prevents catastrophic forgetting by permitting analytic models to adapt to new data streams, which is essential for lifelong monitoring and continuously evolving disease profiles.
- **Semi-Supervised Cross-Modal Learning:** Frameworks such as SPamCo employ co-training, regularization, and pseudo-labeling across multiple data views, effectively leveraging both labeled and unlabeled examples to improve model generalization [36, 63].

These strategies directly confront core limitations in healthcare artificial intelligence, including data sparsity, heterogeneity, and pervasive weak supervision. Integrating noisy or weakly labeled data along with domain expertise into the learning pipeline enables substantial analytic capability without over-reliance on costly expert annotation.

Despite these methodological advances, several unresolved challenges persist. Adaptive fusion models are still susceptible to performance declines under severe distributional shifts that arise due to temporal, contextual, or institutional changes in incoming clinical or IoT data. While federated learning paradigms partially resolve privacy and data localization concerns, they introduce difficulties in aligning model representations across non-i.i.d. datasets and

maintaining the consistency of distributed updates and inferences. Therefore, foundational research remains necessary to harmonize adaptive and privacy-preserving cross-modal learning with robust uncertainty quantification and flexible fusion architectures. Achieving such a synthesis is a prerequisite for realizing the transformational promise of multimodal and IoT-driven analytics in next-generation healthcare systems [36, 49, 63, 78].

## 4 Machine Learning, Deep Learning, and Explainable AI in Healthcare

### 4.1 Model Architectures and Learning Methods

#### 4.1.1 Traditional ML and Deep Learning Architectures in Healthcare.

The landscape of machine learning (ML) and deep learning (DL) architectures in healthcare has evolved considerably, shaped by the increasing diversity and complexity of biomedical data, alongside escalating demands for diagnostic accuracy, scalability, and robustness. Traditional ML algorithms—specifically, Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Decision Trees—have long provided robust baselines for structured data analyses. These algorithms possess notable strengths in settings emphasizing interpretability and computational efficiency, often yielding competitive performance in classification and regression tasks where datasets are moderate in size and efficient feature engineering is feasible [44, 91].

The advent of high-dimensional and heterogeneous biomedical data modalities—including medical imaging, multi-channel physiological signals, and structured electronic health records (EHRs)—has made deep neural architectures indispensable. Among these, Convolutional Neural Networks (CNNs) have become the predominant choice for medical image analysis, excelling at complex feature extraction and hierarchical representation learning in imaging modalities such as X-ray, MRI, and CT [3, 8, 22, 33, 34, 38, 41, 44, 49, 61, 69]. Furthermore, architecture variants such as U-Net and 3D CNNs are tailored for biomedical image segmentation and volumetric delineation, effectively handling spatial complexities inherent in such data [55, 56, 59, 61].

For sequential clinical data, particularly physiological time-series (e.g., ECG, EEG) and EHR data, Recurrent Neural Networks (RNNs)—notably, Long Short-Term Memory (LSTM) networks—have demonstrated substantial efficacy in capturing temporal dependencies [58, 69, 87, 101]. The emergence of Transformer architectures has further augmented the modeling of long-range temporal and contextual dependencies in both time-series and textual modalities. The incorporation of multi-head self-attention within Transformers enables robust performance and flexibility, particularly in complex multimodal fusion scenarios [6, 8, 19, 26, 28, 41, 44, 53, 55, 102].

Emerging models such as Graph Neural Networks (GNNs) are increasingly leveraged to analyze data with intrinsic graph structures, enabling significant advancements in tasks ranging from

molecular property prediction in drug discovery to the modeling of population or knowledge networks in epidemiology [76, 92, 102]. Integrative frameworks combining GNNs and large language models (LLMs) within multi-modal architectures have shown marked improvements in property prediction and knowledge extraction across complex scientific domains [92].

Despite these advances, several persistent challenges limit the full realization of deep learning's promise in healthcare:

- Data annotation scarcity, impeding the training of data-hungry models
- High inter-class variability, which complicates generalization
- Limited generalizability to rare or outlier cases
- Insufficient transparency and interpretability, hindering clinical acceptance

These obstacles remain focal points for ongoing research and methodological innovation [28, 30, 33, 53, 55, 56, 59, 93].

**4.1.2 Transfer, Hybrid/Ensemble, Annotation-Efficient, and Self-Supervised Learning for Multimodal and IoT Healthcare Data.** To address challenges relating to data scarcity and annotation, the field has advanced a variety of methodological strategies, including transfer learning, hybrid/ensemble modeling, annotation-efficient approaches, and self-/semi-supervised learning. Collectively, these methods enhance performance and utility for both unimodal and multimodal healthcare data.

- **Transfer learning** exploits pre-trained models—often established on large, generic, or related datasets—which are subsequently fine-tuned on smaller, domain-specific biomedical datasets. This approach enhances model performance when labeled data is limited or expensive to produce [6, 8, 28, 41, 53, 69, 92]. Frameworks such as COMET use RNNs pre-trained on extensive EHR cohorts and employ early/late fusion of omics and clinical data, achieving improved predictive modeling and deeper biological insight, especially in small-sample contexts [55].
- **Hybrid and ensemble methods**, encompassing strategies like model stacking and combining traditional ML classifiers with contemporary DL architectures, bolster resilience against overfitting and elevate generalizability—an asset in disease diagnosis for oral cancer, cardiovascular disease, and rare event prediction [8, 27, 33, 53, 56, 61, 91]. This integrative approach allows for adaptive model selection driven by contextual priorities, such as balancing accuracy and interpretability.
- **Annotation-efficient paradigms**, such as semi-supervised and self-supervised learning, are critical in environments with limited labeled data. Semi-supervised learning, deploying techniques like pseudo-labeling or consistency regularization, empowers models to exploit vast unlabeled datasets effectively—for example, in prostate MRI segmentation tasks where performance nears that of fully supervised baselines [6, 26, 47].
- **Self-supervised learning** employs pretext tasks or contrastive objectives to extract meaningful representations from unlabeled data. Notably prevalent in transformer-based

and contrastive learning frameworks, these strategies drive superior feature learning for downstream biomedical tasks [6, 30, 51, 69, 78]. In complex settings (e.g., multi-sensor IoT healthcare), self-supervised transformer models and robust pseudo-annotation pipelines, augmented by knowledge distillation, achieve heightened performance and annotation efficiency [6, 8, 26, 28, 53, 55, 78, 79].

Multimodal learning, integrating signals from text, images, structured data, and physiological waveforms, presents considerable promise for health informatics. Multi-modal large language models (MLLMs), equipped with sophisticated fusion and alignment mechanisms such as joint attention networks, have enabled integrated contextual analysis and superior clinical prediction [6, 8, 17, 28, 41, 47, 53, 55, 58, 61, 78]. When combined with transfer learning and self-supervision, these models demonstrate further enhanced capability. However, challenges related to interpretability, computational overhead, and the need for well-aligned multimodal representations persist.

The approaches summarized in Table 4 collectively address major bottlenecks in data efficiency, generalizability, and multimodal integration in healthcare AI.

## 4.2 Explainability, Transparency, and Clinical Trust

**4.2.1 The Imperative for Explainable AI (XAI) Methods and Clinical Applications.** Realizing the transformative potential of ML/DL in healthcare depends critically not only on predictive accuracy but also on model explainability, transparency, and ultimately, the establishment of clinical trust. The imperative for explainable AI (XAI) is pronounced in medical decision-making, where clinicians and regulatory stakeholders require machine learning models to provide not only reliable outputs but also cogent justifications for those outputs [28, 32, 54, 61, 82].

Post-hoc XAI techniques—such as SHAP, LIME, and ELI5—lead the way in quantifying feature contributions and clarifying the sequential reasoning by which models arrive at decisions [28, 44, 61, 64, 82]. These methods have been pivotal in model evaluation for disease diagnostics, including diabetes, monkeypox, and cardiac arrest, where they aid in identifying salient risk factors and producing fine-grained, patient-specific explanations. The integration of XAI tools thus effectively bridges the gap between algorithmic prediction and clinician reasoning [2, 28, 32, 44, 61, 62, 64, 82].

In cutting-edge applications, more advanced explanation modalities are increasingly demanded. Notable examples include:

- **Event-level causal reasoning:** Techniques elucidate causal relationships in clinical events, enhancing diagnostic insight [32, 64].
- **Signal-level interpretability:** Models provide physiologically grounded explanations for quasi-periodic biomedical signals (e.g., ECG, EEG), with methods like SHAP, counterfactual analysis, and event-level causal modeling delivering explanations that align closely with established physiological phenomena and clinical heuristics (for example, arrhythmia detection via R-R interval regularity) [2, 32, 39, 64, 67].

Despite substantial advances, the translation of deep model features into explicitly physiologically or pathologically meaningful entities



Table 4: Overview of Learning Methods and Their Key Contributions in Healthcare

Method	Main Contribution	Key Applications/Examples
Transfer Learning	Leverages pre-trained models to improve performance, especially with limited labeled data	Fine-tuning CNNs for medical imaging [8, 41, 53], RNNs on EHRs (COMET) [55]
Hybrid/Ensemble	Combines diverse models to boost generalizability and reduce overfitting	Disease diagnosis (oral cancer, CVD) [8, 33, 56]
Semi-supervised Learning	Utilizes both labeled and unlabeled data, improving learning efficiency	MRI segmentation [26, 47], pseudo-labeling [6]
Self-supervised Learning	Learns representations via pretext tasks, requiring no labeled data	Transformer-based feature extraction [30, 69, 78]
Multimodal Learning	Fuses diverse data types for richer predictions and context integration	MLLMs with joint attention for clinical text, images, and waveforms [6, 41, 55, 78]

remains a major challenge, underscoring the dynamic and evolving landscape of XAI in healthcare [2, 39].

**4.2.2 User-Centric Analytic Interfaces and Transparency in IoT Healthcare.** Enhancing explainability further, recent progress emphasizes user-centric analytic interfaces that move beyond algorithmic transparency to foster real-time, patient- or case-specific interpretability and direct clinician engagement. Modern analytic platforms have produced self-explanatory interfaces that incorporate predictive modeling with intuitive explanation layers, frequently leveraging SHAP visualizations and tailored for both expert and non-expert audiences [20, 39, 77, 78].

In the domain of IoT-augmented healthcare, system transparency and the capacity for rapid, real-time analytics are paramount for regulatory compliance and clinical adoption [8, 78]. The deployment of transparent analytic layers and dynamic visualization tools not only facilitates trust and uptake but also enables continuous anomaly detection, real-time feedback, and responsive adaptation as IoT systems and clinical environments evolve [20, 39].

**4.2.3 Clinical and Real-World Validation of AI and IoT-Augmented Systems.** The definitive measure of ML/DL and XAI systems’ value lies in their rigorous clinical and real-world validation. This process requires continual evaluation across heterogeneous patient populations, diverse clinical workflows, and variable operating environments. Multicenter studies evaluating AI-augmented radiological segmentation or real-time risk monitoring substantiate that models integrating explainability and user-centric design are more readily adopted and trusted in clinical settings [2, 26, 39, 49, 54, 94].

Despite these successes, robust generalizability remains a pervasive challenge, especially when models confront shifts in data distributions, patient demographics, imaging and scanning protocols, or rare and emergent disease presentations. Key strategies to overcome these limitations include:

- External validation with independent datasets
- Prospective and multicenter studies
- Continuous adaptation and refinement integrated with user feedback

These practices underpin long-term model robustness and ensure ongoing clinical relevance [26, 39, 49].

In summary, the synergy of advanced learning algorithms, annotation-efficient and transfer learning strategies, and the imperative for explainability and transparency encapsulates both the promise and enduring challenges for AI-driven healthcare. The thoughtful and principled integration of these components is essential to realize trustworthy, effective, and equitable machine intelligence in clinical practice.

5 Medical Imaging, Multimodal, and Cross-Modal Analytics

5.1 Automated Segmentation, Registration, and Imaging Diagnostics

Advances in automated medical image analysis have substantially enhanced both the scale and accuracy of diagnostic workflows across such domains as neuroimaging, oncology, and cardiology. Central to this progress are deep learning architectures—such as DenseNet and 3D-UNet—that demonstrate state-of-the-art performance in complex segmentation tasks. These architectures excel by capturing hierarchical and contextual features from high-dimensional imaging data. For example, a two-stage network architecture employing a 2D DenseNet for filtering, followed by a 3D-UNet with integrated dense and residual blocks, achieves robust coronary artery segmentation from CCTA images. This configuration not only yields high Dice Similarity Coefficients but also markedly streamlines preprocessing and computational demands. Furthermore, integrating a Gaussian-weighted merging process enhances segmentation reliability by reducing block-boundary artifacts, underlining the critical role of architectural refinements that propagate both local and global contextual information for anatomical structure delineation [57].

Parallel developments in the registration of digital pathology whole slide images (WSIs) address challenges posed by gigapixel-scale data and staining heterogeneity inherent in histopathological datasets. Recent innovations—such as multi-scale ring encoders—have enabled rapid, automated matching of multi-stained serial WSIs, thereby reducing spatial uncertainties to sub-200  $\mu\text{m}$  levels and significantly accelerating processing relative to manual annotation. These methods facilitate the quantification of tumor heterogeneity and support the optimization of biomarkers like Ki-67, exemplifying the clinical value of methodological advances in enabling personalized cancer diagnosis and treatment [85].

Despite these technical achievements, reproducibility and scalable benchmarking remain significant challenges. The establishment of rigorous multi-center collaborations, introduction of well-characterized reference controls, and the expansion of publicly available annotated datasets are incrementally improving the generalizability of automated diagnostic systems. Nonetheless, large-scale deployment is constrained by annotation scarcity, class imbalance, and marked heterogeneity across imaging protocols and patient populations. These constraints underscore the ongoing need for robust, data-efficient learning strategies and the adoption of standardized evaluation benchmarks [57].

## 5.2 Imaging Data Fusion and Cross-Modal Analysis

**5.2.1 Fusion of Imaging, Biosignals, Laboratory, and IoT/Behavioral Data.** The promise of medical AI increasingly rests upon its ability to synthesize multimodal data—a necessity driven by the multifactorial nature of disease and the inherent heterogeneity of health system inputs. Fusion approaches leverage data from diverse sources, including:

- Imaging modalities (CT, MRI, PET)
- Biosignals (ECG, EEG)
- Laboratory results
- Patient-reported outcomes
- IoT and behavioral sensor data streams

[33, 35, 56, 83].

Multimodal integration not only improves predictive accuracy in disease diagnosis but also bridges the gap between clinical, operational, and environmental data, fostering real-world applicability. Notably, hospital-wide big data platforms capable of integrating clinical, radiological, laboratory, and administrative records in real-time enable both operational efficiencies and novel analytics, supporting large-scale translational research and automated decision support [78].

Methodological generalization beyond medicine—to sectors such as industrial prognosis, urban monitoring, and energy management—has produced valuable design blueprints for structuring, governing, and optimizing heterogeneous data flows [22, 33, 34, 38, 67, 83, 94, 96]. Collectively, these experiences highlight the importance of modular platform architectures, robust data governance models, and high-performance computing backends necessary for translating medical multimodal analytics from research into routine clinical deployment.

**5.2.2 Uncertainty Management, Missing/Correlated Data Handling, and Integrated Diagnostics.** Effective integration of multimodal data streams necessitates comprehensive strategies for managing uncertainty, missingness, and inter-modality correlations—challenges magnified in real-world healthcare settings where data completeness and accuracy are rarely guaranteed. Key methodological advancements include:

- Application of Dempster-Shafer theory to quantitatively manage evidence conflicts, thereby providing comprehensive uncertainty measures.
- Custom loss functions and probabilistic data fusion frameworks that explicitly incorporate both uncertainty and the probabilistic structure of measurement errors [6, 25, 26, 94].
- Data fusion techniques, such as combining molecular embeddings from single-task models, which outperform conventional multi-task models in scenarios characterized by data sparsity or weak intermodality correlations, thus mitigating the impact of data incompleteness and heterogeneity [77].

Furthermore, within patient monitoring and IoT-assisted healthcare, methods such as Proportionate Data Analytics (PDA) enable dynamic classification and adaptive responses to anomalous stream behaviors by disentangling errors from natural physiological variation, thereby enhancing both service response ratios and overall

analytic robustness [8]. Integrative diagnostic frameworks, from machine learning-driven disease risk models to combined radiomics-clinical prognostic tools, exemplify how joint, explicit uncertainty management augments both the trustworthiness and clinical utility of predictive insights.

**5.2.3 Diverse Fusion Strategies (Joint, Late, Early) and Applications.** Operationalizing data fusion in medical domains encompasses a spectrum of strategic models, generally classified into early, joint, or late fusion. The principal characteristics and typical advantages of these strategies are summarized in Table 5.

In healthcare contexts, the decision between early, late, and hybrid fusion depends on data characteristics, the extent and nature of missingness or misalignment, and computational constraints. For example:

- Cross-modal learning frameworks utilizing unbiased retrieval risk estimators and complementary contrastive learning excel even with noisy modality pairings (e.g., mismatched image-text datasets), capitalizing on negative sample regularization to minimize overfitting [47, 51].
- Self-paced multi-view co-training with distributed optimization supports scalable, PAC-learnable integration across multiple heterogeneous views—an adaptation of classical semi-supervised learning suitable for the complexity of modern medical and IoT scenarios [27, 42, 53].

Across these diverse applications, multimodal data fusion is propelling significant advances in patient monitoring, precision diagnostics, and rapid medical multimedia search—catalyzing a broader transition toward holistic, data-driven clinical decision support [8, 28, 30, 31, 33, 34, 36, 47, 61, 63, 77, 87]. Despite ongoing challenges related to generalizability, interpretability, and standardization, methodological innovations are charting a path toward more robust, trustworthy, and fully integrated medical AI. The confluence of advanced segmentation, multimodal data integration, and uncertainty-aware analytics is ultimately shaping the future landscape of precision medicine.

## 6 Operational Analytics, Population Health, and Clinical Deployment

### 6.1 AI-Driven Hospital and Population Health Operations

The convergence of artificial intelligence (AI) and Internet of Things (IoT) technologies is fundamentally transforming operational paradigms in hospitals and population health management, enabling data-driven resource optimization, real-time surveillance, and tailored interventions. In hospital environments, conventional surgery scheduling approaches often result in suboptimal resource utilization and workflow inefficiencies due to insufficient consideration of patient-specific and procedure-specific attributes. With the advent of machine learning, particularly multivariate ensemble models, there has been a marked shift: these systems now harness pre-procedural clinical and administrative data to transcend the limitations of historical mean-based and manually corrected methods. The outcome is a significant reduction in root mean squared error for time predictions and a notable decrease in late-running cases for both elective and acute surgeries. Such advances have immediate

Table 5: Comparison of Fusion Strategies in Multimodal Medical Analytics

Strategy	Integration Stage	Key Advantages	Typical Applications
Early Fusion	Raw data or low-level features combined prior to algorithmic modeling	Enables learning of shared modality representations from the outset; high potential information synergy	Deep neural network training with joint feature embedding
Joint Fusion	Intermediate layers form shared representations during modeling; cross-modal learning possible	Balances cross-modal synergy and robustness; adaptable attention and feature selection	Cross-modal learning (RNNs, Transformers, NetVLAD); video categorization; multimedia retrieval
Late Fusion	Decision outputs of separate models are combined post-modeling	Enhances robustness to missing or noisy modalities; modular adaptation	Ensemble learning; multi-system decision aggregation

clinical ramifications—more precise scheduling directly leads to reduced surgery cancellations, improved operating room throughput, and enhanced predictability in perioperative care processes, collectively contributing to the alleviation of longstanding bottlenecks that negatively impact clinical and economic outcomes [72].

Beyond the operating suite, institution-wide big data platforms are emerging as foundational to the integration of diverse data sources—including clinical, laboratory, and administrative streams—within secure and scalable infrastructures. The West China Hospital Big Data Platform (WCH-BDP) serves as an exemplary model: its architecture is designed for automated, real-time data ingestion and standardized retrieval, harmonizing in excess of 8,000 discrete clinical variables. This robust computational backbone simultaneously advances both operational analytics and AI-driven clinical applications. Critically, the transition from manual to automated data access (reducing retrieval from hours to minutes) underscores the profound potential for such systems to impact not only research productivity but also immediate clinical decision-making, such as critical event detection and real-time management support. These capabilities illustrate the operational benefits of cross-domain, real-time analytics [94].

At the population health level, AI and IoT technologies forge innovative models for epidemiological surveillance and adaptive interventions, yet also highlight enduring challenges. For example, dynamic transmission modeling—combined with detailed economic analysis—demonstrates that focused intervention strategies, such as vector control in high-risk zones, are essential for achieving disease elimination goals in a cost-effective manner. However, these models also demonstrate operational fragility as disease prevalence declines, necessitating ongoing adjustments to surveillance and resource allocation in response to fluctuating program coverage and operational limitations. This underscores the imperative for real-time, adaptive data pipelines that support sustainable epidemic monitoring and the judicious distribution of limited resources, particularly in settings with high variability and resource constraints [47, 52].

IoT-enabled ubiquitous health monitoring illustrates both the potential and the complexity introduced by population-scale analytics. The recently proposed Proportionate Data Analytics (PDA) framework enhances the reliability of anomaly detection within heterogeneous health data streams by dynamically categorizing errors and variations. This approach improves the specificity and responsiveness of digital health services in real-world, error-prone environments by continuously adapting to stream quality and user context. Consequently, PDA represents a critical advancement in robust, quality-aware analytics for scalable remote health monitoring and telemedicine systems [99]. Despite these advances, substantial data-related challenges persist:

- Heterogeneous noise within multi-source data streams
- Incomplete or imbalanced modality representation
- Sample-wise variations in data quality

These factors collectively undermine predictive reliability, especially in settings characterized by multimodal and inconsistent data sources [100]. Effective mitigation of these challenges is essential for the equitable and practical deployment of such analytics at scale.

The interface between digital health platforms and AI analytics is exemplified by large-scale, application-driven interventions in chronic disease management. Deployments integrating continuous glucose monitoring (derived from wearables), dietary and physical activity tracking, and personalized feedback mechanisms have yielded measurable improvements in glycemic control, weight reduction, and behavioral engagement among diabetic and prediabetic cohorts. Notably, these interventions operate autonomously—without dependence on synchronous human coaching—demonstrating the feasibility of scalable population outreach at minimal marginal cost and offering a strategic approach to longstanding gaps in diabetes prevention. However, extending these models across heterogeneous populations necessitates careful consideration of cultural, demographic, and behavioral diversity, alongside stringent long-term validation and thoughtful integration with existing healthcare pathways [5].

Despite rapid advances in digital health capabilities, substantial disparities in care delivery remain entrenched. Analyses of large-scale diagnostic service utilization reveal that, even as technological improvements narrow the access gap, pronounced inequities persist in the timeliness and thoroughness of follow-up care across racial and socioeconomic strata. This observation highlights the necessity of embedding technological innovation within a comprehensive digital equity framework, which addresses not only technical availability but also the accessibility and cultural appropriateness of health interventions [70].

The digitization and interoperability of health information systems have progressed significantly due to targeted national policies and infrastructure investments. Currently, a majority of U.S. hospitals and prescribers possess the capacity for electronic data exchange and public health reporting. Nonetheless, persistent barriers—including technical, usability, and equity-related constraints—particularly impact smaller and rural providers, limiting the fluidity of data necessary for patient-centered, population-level analytics. Overcoming these divides requires continuous enhancement of data standards, development of transparent and trustworthy AI models, and an unwavering commitment to user-centric design principles that facilitate actionable insights at the clinical point of care [10, 94].

6.2 Clinical Decision Support and Human-in-the-Loop Analytics

The steady advancement of operational analytics and population health management is intricately connected to progress in clinical decision support (CDS) systems, where user-centric design and analytic transparency are critical determinants of success. Irrespective of technical sophistication, AI solutions must be intentionally developed for integration with clinician workflows to achieve their

potential in delivering safe, effective, and equitable care. Usability is pivotal in this context: systems that impose cognitive overload, lack transparency in reasoning, or disrupt established workflows are likely to encounter diminished adoption or even active resistance, notwithstanding their superior algorithmic performance [39, 77].

To address these barriers, explainable AI (XAI) frameworks have become indispensable in the pursuit of trustworthy CDS. Methods such as saliency mapping, Shapley Additive Explanations (SHAP), and Local Interpretable Model-Agnostic Explanations (LIME), once confined to academic research, are now operational in clinical contexts. These tools deliver granular, case-specific illumination of the diagnostic logic underpinning complex deep learning and ensemble models, thereby empowering clinicians to understand, scrutinize, and ultimately trust AI-driven decisions. Notable examples include time-series analysis for arrhythmia detection and disease classification in diabetes, where explanatory tools have rendered model outputs accessible and actionable for both clinicians and patients [8, 20].

The integration of analytics with IoT systems further depends on accommodating heterogeneity in data sources and user preferences. In the context of AI-assisted neuroimaging, workflows must support both the delivery of precise quantitative biomarkers and facilitate expert oversight and correction given the notable costs of annotation and the presence of systematic biases in large-scale training datasets. While multimodal fusion remains on the frontier of clinical analytics, the challenges of pervasive data noise, missingness, and intricate interdependencies among modalities demand the development of highly adaptive, transparent, and robust architectures to ensure clinical validity and broad generalizability [8, 100].

At the organizational scale, analytic usability is substantively improved by platforms that balance democratized access to data and models with stringent governance and security. These platforms enable clinicians and operational leaders to derive and act upon real-time insights, supporting the evolution toward self-service, data-driven workflows. Such workflows underpin the human-in-the-loop analytics paradigm, wherein continuous expert feedback calibrates, audits, and refines AI models for local applicability, mitigating risks of model drift and emergent bias [78, 94].

Ultimately, successful CDS deployment requires meticulous alignment with end-user requirements and an explicit focus on clinician trust. Sustainable implementation is attainable only through dynamic, iterative feedback loops between AI developers and clinical stakeholders, ensuring that usability, actionability, and analytic transparency are all practically realized at the point of care. This symbiotic relationship between humans and AI is not simply a matter of user preference; it constitutes a foundational precondition for the delivery of reliable, safe, and ethical analytics in patient care and population health initiatives alike [20, 39, 78].

## 7 Synthetic Data Generation, Privacy, and Security

### 7.1 Generation Approaches and Clinical/IoT Use Cases

The application of synthetic data within healthcare and IoT-driven environments is propelled by the imperative for privacy-preserving

analytics, augmentation of data-scarce scenarios, and the facilitation of regulatory compliance for both clinical and real-time health monitoring contexts [29, 78]. The inherent difficulty in accessing large-scale, real-world patient data—due to legal, proprietary, and privacy constraints—poses significant barriers to effective algorithm development, benchmarking, and deployment across diverse patient populations. Synthetic data provides a pragmatic solution by generating realistic but non-identifiable datasets that enable open exchange, expedite the iterative cycle of hypothesis testing, and support bias analysis. Furthermore, such data can serve as controlled testbeds for operational procedures and regulatory stress-testing [29, 78].

Technologically, the generation of synthetic data for healthcare and IoT applications employs a range of advanced techniques meticulously tailored to replicate the complexity and diversity of medical information. Among the most prevalent frameworks are Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and state-of-the-art natural language processing (NLP) models, which underpin the synthesis of electronic health records (EHRs), medical imaging datasets, and heterogeneous multivariate IoT sensor streams [29]. Agent-based modeling further extends capabilities by emulating intricate temporal patient interactions and disease trajectories—a crucial element for simulating clinical trials or conducting epidemiological studies. The continuous evolution of these methods now encompasses the generation of time-series data, unstructured clinical text, genomic profiles, and cross-modal datasets, aligning with the emerging requirements of precision health monitoring and digital twin paradigms [29, 78].

Key application domains for synthetic data include:

- Privacy-preserving, multicenter AI model training
- Augmentation of data for rare disease cohort analyses
- Simulation of IoT environments for robust edge-device validation
- Policy development and planning at governmental and institutional levels

Collectively, synthetic data is gaining centrality in both centralized and federated analytic workflows—enabling privacy, scalability, and inclusivity across the rapidly evolving digital health ecosystem [29, 78].

### 7.2 Challenges Ethics, Legalities, and Technical Hurdles

Despite the clear benefits, the adoption of synthetic and real-time data streams within healthcare and IoT domains is hindered by a matrix of ethical, legal, and technical challenges that directly impact utility, trust, and societal acceptance. One of the most prominent risks is the perpetuation or even exacerbation of underlying data biases: generative models tuned on incomplete or skewed datasets can yield synthetic records that reinforce inequities or clinical inaccuracies, thus compromising algorithmic fairness and representing a latent threat to marginalized populations [29]. The opacity of many generative processes complicates auditability, making it difficult to ascertain data provenance, validate representativeness, or perform rigorous post-hoc error analyses—functions that are vital for regulatory oversight and quality assurance [78].

Privacy risks persist even with ostensibly de-identified synthetic data. Advances in re-identification techniques—including membership inference and adversarial linkage attacks—raise legitimate concerns that overly realistic synthetic datasets may inadvertently disclose sensitive personal attributes or enable indirect identification of individuals [29, 78]. Existing legal regulations, such as HIPAA (USA) or GDPR (Europe), set baseline requirements for privacy, yet they often lack explicit, harmonized coverage of synthetic data, especially in real-time and streaming IoT contexts. As a result, organizations are frequently left to navigate a patchwork of technical safeguards and jurisdiction-specific compliance requirements [6, 8, 10, 11, 22, 24, 26–30, 33–35, 38, 40, 41, 44–47, 51–53, 58, 60, 61, 64, 73, 75, 79, 89, 90, 92–95, 98–100, 102].

Differential privacy (DP) methods, such as the Private Aggregation of Teacher Ensembles GAN (PATE-GAN), introduce rigorous mathematical guarantees against information leakage; however, factors such as mathematical complexity, tradeoffs between data utility and privacy, and the high computational burden associated with DP can impede widespread adoption [29]. The application of DP in resource-constrained and heterogeneous IoT environments is further challenged by the need for real-time compliance monitoring and dynamic threat mitigation [6, 8, 10, 11, 22, 24, 26–30, 33–35, 38, 40, 41, 44–47, 51–53, 58, 60, 61, 64, 73, 75, 79, 89, 90, 92–95, 98–100, 102].

The regulatory landscape compounds these challenges. Disparities in definitions, enforcement, and coverage across national and supranational levels complicate risk management, while most current frameworks are tailored to static rather than streaming data. This gap exposes real-time IoT systems to unmitigated threats such as sensor spoofing, data poisoning, and unauthorized cross-linkages—vulnerabilities that are not fully addressed by prevailing guidelines [6, 10, 11, 24, 26–30, 35, 41, 46, 47, 51–53, 58, 60, 61, 64, 75, 92–95, 98, 102]. The absence of harmonized definitions and audit mechanisms ultimately limits the establishment of trust among data controllers, practitioners, regulators, and the broader public [29].

A structured comparison of major technical and governance challenges is provided in Table 6, highlighting the interplay and current mitigation limitations.

### 7.3 Advanced Governance and Secure Infrastructures

Overcoming these intertwined challenges requires systemic transformation in the governance, security, and auditability of synthetic, clinical, and IoT-derived datasets. Emerging frameworks emphasize the necessity for end-to-end “digital chain-of-custody” infrastructures, integrating cryptographic elements such as provenance tracking, digital signatures, and audit logs, alongside machine-readable access control policies. These measures collectively enable data accountability from initial generation through to analytical sharing, thereby supporting enforceable regulatory compliance and traceability [1, 6, 7, 11, 12, 14–16, 21, 26–29, 35, 41, 43, 46, 47, 51, 60, 61, 65, 66, 75, 78, 79, 84, 91–93, 95, 98, 102].

The implementation of blockchain and distributed ledger technologies adds a tamper-evident layer to record-keeping, distributes control across multiple stakeholders, and facilitates programmable,

granular access and sharing mechanisms. This is of particular importance in federated and cross-institutional environments, where the boundaries of trust are not always clearly delineated. Distributed ledger-based systems may thereby underpin secure, scalable analysis across globally interconnected IoT networks [1, 6, 7, 11, 12, 14–16, 21, 26–29, 35, 41, 43, 46, 47, 51, 60, 61, 65, 66, 75, 78, 79, 84, 91–93, 95, 98, 102].

In parallel, privacy-preserving analytics—leveraging cryptographic primitives such as homomorphic encryption, secure multi-party computation, and privacy-aware machine learning—enables computation on encrypted or masked data, thereby ensuring that even data processors do not gain direct access to sensitive values. Within IoT settings, the adoption of edge intelligence and decentralized architectures further reduces exposure risk by confining sensitive computation and storage to device-local or micro-environmental domains [1, 6, 7, 11, 12, 14, 15, 21, 26–29, 43, 46, 47, 51, 60, 66, 78, 84, 91, 95, 98].

Nevertheless, robust governance remains a continuously evolving frontier that demands alignment among technical innovations, regulatory adaptations, and ethical imperatives, including transparency and equity. Sustained progress is contingent on interdisciplinary collaboration among technologists, clinicians, policymakers, ethicists, and patient groups in order to define best practices, monitor emergent risks, and continuously refine adaptive, context-sensitive controls that evolve in parallel with the technology itself [29, 78].

## 8 Data Quality, Benchmarking, and Technical Robustness

### 8.1 Data Quality in Clinical and IoT/Streaming Contexts

*8.1.1 Problems of Data Heterogeneity, Class Imbalance, and Missing Modalities in Biomedical and IoT Data.* Emerging applications of artificial intelligence in healthcare, biomedicine, and the Internet of Things (IoT) are increasingly reliant on the availability and reliability of heterogeneous, multimodal data. Biomedical datasets are intrinsically diverse, encompassing structured clinical records, medical imaging, physiological signals, and sensor streams from wearables or IoT devices. Each of these data sources exhibits distinct sampling rates, noise profiles, and patterns of data completeness. This inherent heterogeneity poses significant challenges for data integration and model generalizability, impeding robust clinical translation. Models developed on homogeneous or single-modality datasets frequently fail to maintain performance when confronted with the complexities of real-world, multimodal environments [10, 24, 100].

A particularly persistent issue within these domains is class imbalance. In both clinical and streaming contexts, events of greatest concern—such as seizures or disease onsets—are typically rare compared to the abundance of normal observations. This imbalance can bias predictive models toward the majority class, thus diminishing sensitivity to rare but clinically critical outcomes [24, 94]. Compounding these challenges, missing data and incomplete modalities—arising from technical failures, selective recording, or inconsistent patient participation—further compromise data utility. IoT-enabled healthcare systems are especially vulnerable, with streaming data subject to loss, corruption, asynchronous arrival, and device

**Table 6: Overview of Major Challenges in Synthetic Data for Healthcare and IoT**

Challenge Category	Manifestation in Synthetic Data Contexts	Current Limitations of Mitigation Strategies
Bias and Fairness	Propagation or amplification of demographic and clinical inaccuracies	Limited model transparency; insufficient bias auditing tools; difficulty representing rare subpopulations
Auditability and Provenance	Opaque generative models obscure source-to-sample tracking and error traceability	Lack of standardized audit frameworks; complexity of generative architectures restricts interpretability
Privacy and Re-identification	Vulnerability to membership inference and linkage attacks; residual sensitivity in synthetic samples	Incomplete legal coverage; computational and utility tradeoffs with privacy-enhancing technologies
Legal and Regulatory Gaps	Inconsistent definitions across jurisdictions; static data-focused frameworks overlook streaming/IoT realities	Fragmented regulatory guidance; lack of harmonized standards; limited enforcement for IoT applications

heterogeneity, all of which complicate both upstream preprocessing and downstream inference [8, 78].

These data quality challenges often interact in complex and compounding ways. Data heterogeneity, class imbalance, and missingness rarely occur in isolation; their intersection can produce intricate failure modes. For example, the loss of even a single modality in a data fusion task can irreversibly confound the predictive pipeline unless robust imputation or modality-aware strategies are employed. Empirical investigations consistently identify these factors as primary obstacles to the clinical deployment and sustained, real-world operation of machine learning systems in biomedical and IoT domains [78, 94, 100].

**8.1.2 Technical Preprocessing: SMOTE, Artifact Rejection, Denoising, and Anomaly Detection for IoT and Wearable Streams.** Multiple sophisticated preprocessing pipelines have been developed to address the challenges posed by data quality deficiencies in biomedical and streaming contexts. For combating class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) and related algorithms have become standard. By generating synthetic samples for underrepresented classes, SMOTE can partially correct skewed class distributions, enhancing detection of rare yet clinically important events [37, 77]. However, SMOTE’s utility is limited with highly nonstationary stream data or in ultra-high dimensional settings, where it may exacerbate overfitting or introduce artificial boundaries if applied naively [78].

Artifact rejection and signal denoising constitute fundamental components of preprocessing for physiological signals, especially those collected in uncontrolled, real-world environments. Traditional methodologies employ filtering, thresholding, and dimensionality reduction via principal component analysis. More recently, deep learning approaches—including convolutional autoencoders and dual-stream neural networks—have shown promise in distinguishing complex, context-dependent artifact signatures from true physiological signals [71, 80]. In the realm of IoT and continuous streaming, anomaly detection techniques such as Proportionate Data Analytics (PDA) facilitate real-time segmentation of data, distinguishing between normal patterns and artifacts or device malfunctions [8].

Despite technical advancements, balancing denoising selectivity against preservation of clinically relevant information remains a central challenge. Excessive denoising may eliminate subtle but significant signals, whereas insufficient artifact rejection leaves residual confounders. Emerging approaches increasingly integrate artifact-aware loss functions and explicitly model uncertainty during both preprocessing and downstream training, supporting more nuanced management of these trade-offs [78].

As summarized in Table 7, each preprocessing technique exhibits context-dependent strengths and weaknesses, underscoring the

need for tailored solutions and integration of uncertainty modeling in pipeline design.

## 8.2 Benchmarking and Validation

**8.2.1 Multicenter, Real-World, and Robust Benchmarking Approaches.** Rigorous benchmarking is fundamental to the development and validation of generalizable machine learning models in biomedicine and IoT. While single-center, retrospective studies have historically been prevalent, the field is shifting toward multicenter and real-world validation paradigms. This evolution is motivated by recurring failures of models that perform well under limited, homogeneous conditions but falter upon deployment in diverse clinical environments [1, 41, 92]. Initiatives such as the Image Biomarker Standardization Initiative (IBSI) exemplify the critical importance of harmonizing feature definitions and preprocessing protocols, thereby enabling reproducibility, transparent reporting, and robust validation across different software platforms [53, 55, 61].

Beyond imaging, the creation of large-scale, multicenter data platforms that encompass multimodal clinical and streaming information has become instrumental for assessing algorithmic generalizability [19, 24, 28, 75, 84, 88, 102]. These endeavors foster new benchmark datasets that incorporate real-world complexities—including missing data, fluctuating class distributions, and heterogeneous device characteristics—facilitating stress-testing of algorithms before clinical or operational deployment [24]. Contemporary benchmarking protocols now emphasize open-source curation, transparent metrics, and prospective, temporally-anchored validation, ensuring that reported performance is not merely an artifact of retrospective overfitting [1, 41, 53].

**8.2.2 Addressing Computational/Deployment Limitations, Dataset Imbalance, and Patient Variability.** In practice, benchmarking exposes a convergence of technical and systemic obstacles that must be overcome to enable scalable AI deployment. Computational limitations persist, as deep learning models—despite their strong performance on curated datasets—often require resources exceeding those available in typical clinical or edge computing environments. Real-time analysis of high-resolution streaming data is thus frequently infeasible without model optimization. To address these constraints, research increasingly focuses on neural network quantization and the development of computationally efficient architectures, which are essential for practical edge deployment in both clinical and IoT settings [2, 18, 37, 78].

In addition, data imbalance and patient variability introduce further challenges. Achieving truly robust benchmarking requires evaluations that are stratified across demographic, clinical, and device-specific subgroups. This includes explicit quantification of performance disparities and the application of non-inferiority analyses for underrepresented populations [77, 78]. Addressing these challenges involves:

Table 7: Comparison of Preprocessing Techniques for Biomedical and IoT Data

Technique	Primary Target	Key Advantages	Limitations
SMOTE	Class imbalance	Supports rare event detection, widely adopted in tabular data	Limited efficacy for high-dimensional, nonstationary streams; risk of artificial boundaries
Filtering/Thresholding	Signal artifacts, noise (structured signals)	Simple, fast, interpretable	May remove significant signal components; less effective with complex artifacts
Deep Learning Denoising (Autoencoders, dual-stream nets)	Complex, context-dependent noise	Adaptability to complex artifacts; automated feature extraction	High computational demand; requires large labeled datasets
Anomaly Detection (PDA)	Device failure, data corruption (streams)	Real-time operation, models error characteristics	May miss subtle, context-specific anomalies; depends on accurate error modeling

- Algorithmic innovations (e.g., domain adaptation, meta-learning, sample reweighting)
- Infrastructure improvements for federated, multi-institutional data aggregation
- Statistical methodologies for subgroup-specific performance assessment [18, 37]

**8.2.3 Ensuring Translation to Scalable and Interoperable Platforms.** Even when technical performance metrics are met, the broader translational impact of AI systems in healthcare and IoT depends fundamentally on integration within scalable and interoperable platforms. The proliferation of heterogeneous health IT infrastructures has led to fragmented data ecosystems, impeding efficient data exchange, aggregation, and systematic benchmarking [28, 35, 92, 102]. Recent progress—including advances in interoperability standards, terminology harmonization, and development of federated system architectures—has begun to address these silos, supporting more seamless deployment of AI solutions across distributed settings [35, 97, 102]. Persisting barriers, however, comprise non-standardized data formats, privacy and compliance constraints, heterogeneous computational capacity, and limited adoption of open APIs. Overcoming these hurdles is essential for realizing robust, scalable, and equitable AI-enabled healthcare and IoT ecosystems [28, 75].

9 Key Challenges, Open Problems, and Future Directions

9.1 Technical, Methodological, and Practical Challenges

The deployment of advanced analytics and artificial intelligence (AI) in healthcare, particularly within environments leveraging the Internet of Things (IoT), real-time monitoring, and resource-constrained contexts, is limited by a confluence of technical and practical barriers. Chief among these is data heterogeneity—the presence of diverse data sources, modalities, formats, and qualities originating from clinical, sensor-based, and operational systems. This heterogeneity complicates data integration pipelines, frequently leading to datasets characterized by incompleteness, inconsistency, or noise, which undermines algorithmic robustness and generalizability [2–4, 6, 8, 10, 11, 18, 22, 24, 25, 28, 33–38, 41–44, 47–49, 53, 55, 56, 59, 61, 63, 64, 67, 69, 75–79, 83, 84, 86, 88, 91–94, 96, 100, 101].

The scarcity of high-quality annotations compounds these issues. Many AI and deep learning models require extensive labeled datasets, yet such curated resources remain limited—especially for rare diseases, geographically diverse cohorts, and novel data sources such as wearable biosensor streams or molecular diagnostics. The challenge is particularly acute in medical imaging, where annotation relies on scarce human expertise and labor-intensive manual

processes, creating bottlenecks for both model training and external validation [28, 53, 56, 69, 96].

These barriers are magnified when scaling analytic solutions, particularly for real-time or edge deployment in resource-limited settings such as small clinics or rural centers with constrained infrastructure. State-of-the-art models—such as large multimodal architectures or federated learning systems—often exceed the computational and memory capacities of IoT devices, which are limited in power, storage, and connectivity [3, 6, 8, 22, 24, 25, 33–35, 38, 47, 53, 76, 78, 88, 91, 100]. Additional operational challenges include data synchronization issues, streaming errors, packet loss, latency, and device failures, all of which require robust methods for anomaly detection, missing-data imputation, and adaptive model retraining. Recent work has focused on unsupervised and semi-supervised learning to reduce annotation demands, edge-optimized model architectures, and hardware-aware model pruning [22, 37, 44, 47, 53, 59, 79, 96, 100]. Nonetheless, standardization and broad implementation of these solutions remain elusive, highlighting a persistent scalability and practicality gap for real-world applications.

9.2 Interpretability and Clinical Impact

The effective translation of analytic innovations into clinical practice is contingent upon interpretability, clinical utility, and generalizability. High-performing, yet opaque, models—particularly deep neural networks and large foundation models operating on images, signals, and clinical text—often function as “black boxes.” This opacity undermines clinician trust and impedes both validation and regulatory acceptance [2, 8, 20, 22, 25, 28, 32, 34, 35, 39, 53, 54, 56, 61, 62, 64, 67, 78, 82, 100]. Clinical utility is intimately bound to interpretability: beyond providing predictions (such as disease classification or risk scores), models must articulate the underlying rationale, delineating salient image areas or critical features in time-series data. Recent advances in explainable AI (XAI)—including post-hoc saliency mapping, attention visualization, and Shapley value-based feature attribution—enable both global and case-specific interpretative insights, facilitating scientific validation and regulatory review [2, 22, 39, 53, 54, 64, 82, 100].

A distinct but interrelated concern lies in ensuring analytic generalizability across diverse clinical settings, populations, and data acquisition systems. Many models are prone to overfitting to specific institutions, cohorts, or medical devices, resulting in decreased performance in cross-site or prospective validations [2, 25, 28, 44, 56, 62, 78]. This phenomenon is observed across domains, from radiomics to real-time vital sign prediction. Efforts to enhance generalizability include domain adaptation, transfer learning, multi-site federated learning, and incorporation of biological or clinical priors into model regularization. Nevertheless, rigorous external validation and systematic evaluation at scale remain insufficient [2, 25, 34, 39, 44, 54, 78, 100].

### 9.3 Standardization, Equity, and Ethical Considerations

Advances in analytic healthcare depend upon multifaceted standardization and an unwavering commitment to equity and ethical principles. Across modalities—including imaging, -omics, and electronic health record (EHR) data—variability in workflows, acquisition protocols, annotation definitions, metadata schemas, and feature nomenclature hinders reproducibility and impedes model transferability [1, 1, 3, 3, 6, 8, 22, 27, 27, 28, 28, 33, 35, 35, 38, 38, 42, 46, 47, 47, 51, 51–53, 53, 54, 56, 61, 61, 69, 69, 78, 83, 84, 94, 96, 99, 100, 100]. International initiatives, such as the Image Biomarker Standardization Initiative (IBSI), have advanced harmonization in specific domains like radiomics. However, non-standardized practices in feature extraction, segmentation, and reporting continue to undermine multi-center reproducibility and clinical translatability [3, 8, 28, 35, 38, 53, 69, 78, 100].

Equity in analytic healthcare extends beyond algorithmic fairness to encompass broader social determinants and the digital divide. Persistent disparities are shaped by infrastructural deficits, inconsistent connectivity, device interoperability limitations, and the geographic maldistribution of expertise and support [5, 8, 10, 28, 47, 52, 70, 78, 94, 99, 100]. To promote both reproducibility and fairness:

- Clinical workflows and analytic protocols must be harmonized across domains and locations.
- Bias mitigation, transparent reporting, and inclusive dataset curation are essential.
- Investments in equitable infrastructure and participatory design should be prioritized, particularly for marginalized or underserved populations.

These steps aim to mitigate model bias and advance equitable patient outcomes across demographic and resource strata [1, 3, 8, 28, 38, 42, 47, 51, 54, 61, 69, 78, 83, 84, 94, 96, 99, 100]. In the context of increasing reliance on digital health solutions for chronic disease management and pandemic response, these concerns become even more critical.

### 9.4 Privacy, Security, and Compliance

As analytic and data-sharing platforms become increasingly sophisticated, ensuring privacy, security, and regulatory compliance grows ever more imperative. Legislation such as the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) mandates stringent controls over data sharing, necessitating advanced technical solutions for privacy-preserving analytics and secure pipeline design [6, 8, 10, 11, 22, 24, 26–30, 33–35, 38, 40, 41, 44–47, 51–53, 58, 60, 61, 64, 73, 75, 78, 79, 89, 90, 92–95, 98–100, 102]. Emerging threats encompass data breaches, risks of re-identification, and the use of synthetic data for privacy enhancement—all of which introduce new attack surfaces and ethical dilemmas. Techniques such as federated learning and differential privacy are central to technical mitigation; however, they introduce computational burdens and auditability challenges that are yet to be fully resolved [8, 22, 26, 28, 29, 33, 38, 45, 46, 53, 78, 89, 94, 99, 100].

Compounding these considerations is the dynamic and continuous nature of data originating from wearables, IoT devices, and

streaming EHR feeds. This necessitates adaptable legal, technical, and ethical frameworks to ensure auditability and accountability as both data and clinical applications evolve [6, 8, 10, 22, 28, 29, 34, 38, 47, 52, 59, 78, 79, 94, 100]. Among the foremost challenges are:

- Establishing scalable governance for data access and secondary use
- Designing fair and comprehensible consent models suitable for ongoing data collection
- Conducting regular audits, especially in multi-center and cross-border collaborations
- Integrating privacy-by-design and transparency-enabling toolkits from the outset

Despite growing recognition, consistent protocols for compliance verification and accountability remain underdeveloped.

### 9.5 Algorithmic Innovation, Architecture, and Recommendations

Recent algorithmic progress has led to the emergence of continual learning, federated learning, hybrid joint models, and self-supervised approaches, all geared toward scalability and adaptability in healthcare environments infused with IoT devices and intermittent connectivity [3, 6, 8, 22, 26, 28, 33–36, 38, 47, 49, 53, 58, 61, 63, 64, 87, 88, 99, 100]. These methods are designed to address data heterogeneity, reduce the need for costly annotation, and enhance privacy by facilitating learning from decentralized datasets. Furthermore, there is increasing advocacy for architectures that promote modularity, plug-and-play interoperability, and adherence to ethical design principles, thereby accelerating deployment and alignment with evolving clinical and regulatory standards [1, 3, 6, 7, 11, 12, 14–16, 21, 22, 25–28, 30, 33–35, 38, 42–44, 47, 49, 51, 52, 54, 55, 58, 60–62, 65, 66, 74, 75, 78, 82, 84, 89, 91–94, 99, 100, 102].

However, implementation faces several persistent tensions, as outlined in Table 8, which illustrates the inherent trade-offs shaping adoption and ongoing research.

A related concern is the paucity of validated design frameworks and benchmarks for modular analytic platforms that ensure broad interoperability, ethics, and transparent governance across clinical domains. Promising initiatives focusing on open data standards, API-driven integration, modular software architecture, and responsible AI practices demonstrate potential, though require more extensive adoption and cross-sectoral collaboration [1, 3, 6, 22, 25, 26, 28, 33–35, 38, 47, 49, 51, 52, 54, 55, 60, 61, 78, 84, 89, 93, 94, 100]. Future architectures must explicitly embed auditability, explainability, scalability, and security, while accommodating requirements unique to genomics, imaging, clinical text, and temporal data streams.

### 9.6 Summary

In summary, despite significant advances across the spectrum of healthcare analytics, persistent challenges—spanning data heterogeneity, annotation scarcity, system scalability, interpretability, equity, standardization, privacy, compliance, and architectural sustainability—remain central to both research and operational practice. Integrative approaches that harmonize technical rigor, transparency, and ethical foresight are essential to realizing the transformative potential of analytics and AI in medicine.



**Table 8: Core Algorithmic Tensions Limiting Healthcare Analytics Deployment**

Aspect	Desirable Outcomes	Limiting Trade-offs/Challenges
Explainability	Transparent, trustworthy predictions; regulatory compliance	Reduced model accuracy, higher computational cost
Interoperability	Seamless integration with legacy and modern systems	Increased system complexity; standardization gaps
Resource Efficiency	Feasibility on edge/IoT devices; sustainable operation	Model miniaturization can reduce accuracy; limited scalability
Ethical Governance	Responsiveness to legal, clinical, and societal requirements	Fragmented standards; limited auditability and benchmarking

## 10 Scalability, Toolkits, Standards, and Analytical Ecosystems

### 10.1 Key Datasets and Software Toolkits

The exponential escalation of multi-modal, clinical, and Internet of Things (IoT) data volumes has necessitated a robust foundation of centralized datasets and open-source software toolkits. These developments underpin both methodological advances and translational applications across healthcare and analytical domains. Among the most influential centralized datasets is UniMod1K, which offers synchronized RGB, depth, and language data—addressing the long-standing challenge of multi-modality alignment—thus supporting advanced fusion models for object tracking and monocular depth estimation [85]. ImageNet-ESC, extending the paradigmatic ImageNet dataset into the audio-visual sphere, further promotes few-shot learning by leveraging cross-modal associations and is extensively employed for benchmarking adaptation capabilities in both academic and industrial settings [55, 83].

In digital health, the MIT-BIH Arrhythmia database and CHB-MIT seizure corpus provide vast, annotated electrocardiogram (ECG) and electroencephalogram (EEG) datasets. These resources are invaluable for developing and validating machine learning algorithms, fostering reproducible studies in arrhythmia detection and seizure prediction, respectively [3, 8, 13, 17, 19, 37, 38, 48, 58, 77–79, 86, 96]. Kaggle’s intracranial EEG datasets and associated challenge environments expedite research in personalized and real-world seizure forecasting by providing high-resolution, longitudinal brain recordings, thereby bridging the divide between controlled laboratory settings and clinical practice [78, 79]. Collectively, this data ecosystem facilitates algorithmic benchmarking, enables rigorous cross-institutional comparisons, and underpins the progress of scalable, generalizable AI in healthcare and sensor-driven analytics.

Software toolkits have concomitantly evolved to address the growing heterogeneity and complexity of analytics challenges. The `scikit-multimodallearn` library exemplifies general-purpose, open-source design tailored for multi-modal supervised learning, offering seamless integration with the broader Python machine learning ecosystem [18]. Support for flexible data formatting and modality-specific workflows makes it a preferred choice for rapid development and prototyping. Radiomics platforms—particularly those stemming from the Image Biomarker Standardization Initiative (IBSI)—play a pivotal role in standardizing the extraction and interpretation of high-dimensional features from medical images across modalities such as CT, PET, and MRI, thereby substantially improving reproducibility and diagnostic reliability [13, 36, 63]. The increasing prevalence of open-source code repositories associated with recent publications has further democratized access

to cutting-edge techniques in cross-modal analytics, sensor integration, medical image segmentation, and multi-faceted behavioral analysis [49, 78]. Notably, these repositories often supply pre-trained models, annotation utilities, and reproducible computational pipelines, thus lowering barriers to clinical and industrial adoption. The growing emphasis on modularity, scalability, and thorough documentation (supported by comprehensive application programming interfaces) testifies to the maturation of the software ecosystem, aligning directly with the technical demands of large-scale analytical deployments.

Beyond data and software, the rigor of annotation and evaluation remains foundational for robust analytics—particularly in high-stakes domains such as clinical validation and IoT-driven automation. Standardized metrics, including positive percentage agreement (PPA), positive predictive value (PPV), response and delivery times, variation factors, and delivery accuracy, are now routinely utilized to enable fair benchmarking and to monitor performance against established baselines [1, 27, 28, 35, 42, 47, 51, 53, 54, 61, 76, 78, 84, 85, 99]. These metrics are embedded within regulatory frameworks, for example, those governing next-generation sequencing and imaging analytics, where stringent, error-based workflows calibrated to reference datasets are required for reliability, reproducibility, and auditability [53, 61]. Nevertheless, persistent challenges such as annotation format heterogeneity, class imbalance, inconsistencies in ground truth assignment, and domain-specific metric selection can undermine generalizability and translational capacity if not rigorously addressed [27, 28, 47, 51, 78, 99]. Continued attention to these issues is essential as the field advances toward automated, interoperable analytics and clinical decision support systems.

### 10.2 Secure, Scalable, and Interoperable Analytical Ecosystems

The expansive proliferation of distributed data sources—including federated clinical repositories and diverse IoT sensor streams—has crystallized the demand for analytic ecosystems that are secure, privacy-preserving, interoperable, and efficient at scale. Distributed and federated analytic approaches have emerged as primary solutions to the privacy constraints inherent in health and IoT data. These approaches enable collaborative model training without necessitating cross-institutional raw data exchanges [7, 11, 12, 14–16, 21, 41, 43, 46, 60, 65, 66, 75, 91–93, 95, 98, 102]. For instance, federated learning strategies such as FedAvg and FedProx have demonstrated near-centralized performance in distributed medical imaging tasks, all while maintaining robust privacy guarantees [41]. These frameworks have been evaluated on heterogeneous datasets across multiple modalities and pathology domains,

highlighting strong generalizability. However, they also reveal persistent limitations, including inter-site data heterogeneity, communication bottlenecks, and challenges synchronizing model updates [10, 46, 75, 94, 95].

Security and auditability are increasingly enhanced by the integration of blockchain and related ledger technologies, which offer transparent, tamper-evident transaction records of both data access and analytic workflow execution [7, 12, 14, 16, 21, 41, 62, 65, 91, 93]. Combining blockchain-backed data provenance, secure model aggregation, and robust API-level integration enables the construction of verifiable analytical pipelines, which are crucial for regulated clinical environments and IoT deployments alike [1, 6, 26–28, 35, 47, 51, 61, 66, 78, 79, 84, 102]. The merger of federated analytics and blockchain architectures addresses pressing concerns such as data sovereignty, tamper resistance, and multi-institutional coordination. Nevertheless, these benefits often come at the cost of increased system complexity, transaction latency, and heightened resource demands, necessitating careful trade-off analysis in real-world settings [1, 8, 78].

The assurance of efficiency, reproducibility, and regulatory compliance in large-scale, interconnected deployments demands not only technical innovation but also robust standardization and governance frameworks [22, 26–28, 30, 33, 34, 38, 47, 49, 51, 53, 54, 58, 78, 79, 92, 94, 95, 99]. Landmark initiatives—for example, the IBSI project for radiomic feature harmonization and major efforts toward electronic health record interoperability—have yielded essential consensus protocols guiding data formatting, algorithmic benchmarking, and metric reporting, thereby fostering reproducibility and translational potential [13, 30, 36]. Comprehensive data platforms such as WCH-BDP exemplify these principles through their integration of interoperable data lakes, rigorous governance with standardized terminologies, natural language processing-based data structuring, and layered security frameworks, all designed to enable rapid and secure analytics at scale [64]. Notwithstanding these advances, the reality of deployment continues to be shaped by a variety of ongoing challenges:

- Incomplete or imperfect standardization across data sources and analytic processes
- Persistent data silos and mismatched ontologies hindering integration
- Technical heterogeneity across institutions limiting interoperability
- Disparities in resource allocation impacting scalability
- Evolving requirements for multi-stakeholder governance and consensus

Superimposed on these infrastructure issues is a pressing requirement for analytical transparency, explainability, and trustworthiness, prerequisites for both clinical adoption and regulatory acceptance. Techniques such as comprehensive model auditability, embedded conceptual knowledge representation, graph-based causal modeling, formal verification, and explainability have been identified as foundational for future-proofing analytic pipelines and ensuring compliance within medical AI and IoT ecosystems [49, 53]. Despite their critical importance, these approaches frequently encounter real-world barriers: high annotation overhead, complexity

of deployment, and challenges to computational scalability all temper their immediate impact in operational settings. Addressing these limitations requires sustained collaborative efforts spanning technical, clinical, and governance domains, including embedded standards and co-design with end-users. Such efforts remain pivotal for building impactful, sustainable analytics infrastructures across health, scientific, and industrial sectors.

## 11 Conclusions, Synthesis, and Outlook

### 11.1 Synthesis of Advances Across Biomedical Signal Processing, Multimodal and Cross-Modal Analytics, Deep Learning, IoT/Real-Time Systems, Synthetic Data, and Explainable AI

The rapid convergence of advances in biomedical signal processing, multimodal and cross-modal analytics, deep learning, IoT-enabled real-time monitoring, synthetic data generation, and explainable artificial intelligence (XAI) has fundamentally reshaped the landscape of healthcare analytics and clinical operations [1–15, 17–39, 41–44, 47–49, 51–64, 66–72, 75–81, 83–88, 91–94, 96, 97, 99–102]. Signal processing remains the methodological foundation for precise feature extraction from biosignals such as EEG, ECG, sEMG, and medical imaging. Methods including rTV-gPDC and multiresolution decomposition are now routinely employed to improve signal integrity, attenuate noise and artifacts, and address nonlinearity within both clinical and IoT-driven contexts [15, 21, 91].

In parallel, deep learning architectures—particularly convolutional neural networks (CNNs) and transformers—have established state-of-the-art benchmarks in medical data classification, detection, segmentation, and registration, in many cases achieving diagnostic performance on par with, or exceeding, human experts [1, 13, 42, 52, 53, 55, 58, 76]. The effectiveness of these models is significantly enhanced by sophisticated innovations in data fusion and multimodal learning. Frameworks leveraging cooperative, late, early, and intermediate joint representation learning, as well as cross-modal adaptation, efficiently resolve challenges associated with data heterogeneity and complementarity, thereby enabling robust domain adaptation and comprehensive modeling of physiological and regulatory processes [19, 28, 30, 32, 41, 51, 62, 64, 75, 79, 88, 92, 93, 97, 102].

The advances described above are accompanied by a transformation of healthcare delivery, largely attributable to the integration of IoT and real-time systems. Pervasive sensor networks, scalable analytics pipelines, and federated computation now support point-of-care diagnostics, remote home monitoring (e.g., wearable EEG/IMU systems for epilepsy detection), and the optimization of operational workflows spanning entire healthcare systems [49, 50, 60, 78]. Key methodological progress further stems from the introduction of privacy-preserving synthetic data generation techniques, such as generative adversarial networks (GANs), variational autoencoders (VAEs), and agent-based simulations. While these approaches address limitations in data sharing, rare-event modeling, and regulatory compliance, they simultaneously necessitate the development of robust digital trust and audit frameworks to mitigate risks relating to bias, re-identification, and provenance tracking [57, 79, 81].

Additionally, the embedding of explainable AI mechanisms—most notably, SHapley Additive exPlanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME), and post-hoc saliency mapping—directly facilitates regulatory transparency, clinician trust, and eventual clinical adoption. Empirical evidence demonstrates tangible benefits of these mechanisms in disease diagnosis, risk prediction, and operational decision support settings [6, 20, 29, 75, 79]. Collectively, these advances are propelling a transformation from reactive, retrospective healthcare toward a proactive, data-driven, and adaptive paradigm.

## 11.2 Recap of IoT-Specific Strategies: Proportionate Data Analytics, HRLS-TS for Real-Time Monitoring, and Dynamic Error Detection

IoT-focused data analytics methodologies have matured considerably, enabling them to address the inherent complexity, noise, and heterogeneity of biomedical and operational data streams. Proportionate Data Analytics (PDA) is emblematic of this progress, providing an adaptive framework that differentiates between natural variation and genuine anomalies within heterogeneous real-time signals. This discrimination is essential for maintaining a balance between sensitivity and specificity, preventing system overload and reducing the incidence of false alarms [78]. PDA achieves these objectives by dynamically calibrating error detection thresholds through statistical modeling—particularly regression-based error/variation decomposition across multiple time windows. As a result, it sustains high identification and delivery ratios, even under variable data quality.

Furthermore, the Hybrid Recurrent Long Short-term based Tyrannosaurus Search (HRLS-TS) algorithm illustrates the potential of synergistically combining deep learning and evolutionary computation for real-time, context-aware biomedical monitoring. For instance, in the domain of menstrual physiology tracking, HRLS-TS integrates LSTM-based temporal modeling with adaptive metaheuristic optimization for hyperparameter and feature selection. This composite approach achieves both high predictive accuracy and computational efficiency in nonstationary, streaming IoT healthcare scenarios [49]. Collectively, innovations such as PDA and HRLS-TS typify a broader evolution in IoT healthcare analytics: they reflect the necessity to accommodate signal variability, robustness to noise, and real-world scalability, while producing actionable insights rapidly for clinical and operational utility.

## 11.3 Implications for Disease Detection, Personalized/Precision Medicine, and Clinical/Operational Workflow Integration

The convergence of technological and methodological breakthroughs discussed herein is accelerating healthcare transformation across disease detection, personalized and precision medicine, and workflow optimization. For disease detection, proactive analytics leveraging multimodal and cross-modal signals—including organ-specific liquid biopsies, multi-omic profiles, and continuous sensor-based physiological monitoring—facilitate earlier risk identification, often preceding the onset of overt symptoms or irreversible tissue

pathology [10, 14, 52, 57, 79]. In key domains such as infectious disease and oncology, rapid, scalable, and high-resolution stratification is now achievable—propelled by advances in molecular diagnostics, digital pathology, and AI-driven imaging and sequence analysis [9, 11, 27, 30, 43, 52, 56, 63, 79, 83, 96].

Personalized and precision medicine initiatives are increasingly feasible as granular, multimodal patient phenotyping becomes the norm; integration across clinical, molecular, imaging, sensor, and environmental data streams enables sophisticated analytics pipelines capable of modeling intricate interactions and inferring causal relationships [23, 28, 36, 41, 42, 44, 52, 53, 55, 57, 92, 101, 102]. Research demonstrates that real-time digital interventions—including continuous glucose monitoring feedback systems and multi-source longitudinal tracking—lead to more tailored treatments, dynamic intervention adjustments, improved adherence, and measurable outcome improvements for chronic and complex conditions [8, 50, 81, 86].

Clinical and operational integration is similarly being redefined. Automated systems for segmentation, scheduling, triage, and capacity management—enabled by deep learning, time-series analytics, and explainable AI—reduce diagnostic bottlenecks, optimize resource allocation, and minimize delays in critical care pathways (e.g., cardiology, surgery, radiology) [3, 6, 22, 25, 38, 55, 85, 86]. In parallel, federated and cloud-based hospital data platforms are driving the interoperability and accessibility of analytics across multi-institutional settings, though continuing challenges in data harmonization and standardization persist [17, 77]. The ongoing adoption of explainable AI and integrated feedback systems is laying the groundwork for a more transparent, collaborative, and continuously learning healthcare environment.

## 11.4 Future and Persistent Challenges: Standardization, Global Scalability, Heterogeneity, Explainability, Digital Equity, Privacy/Security, Responsible AI, and Collaborative Research

Despite significant progress, several persistent and systemic challenges must be addressed to translate these advances into large-scale, equitable, and clinically impactful solutions. Chief among these is the need for rigorous standardization—both in feature definition (for example, through initiatives such as the Image Biomarker Standardization Initiative and advances in radiomics/texture analysis) and in clinical reporting. Even subtle inconsistencies in data computation or documentation can undermine reproducibility, limit interoperability, and impede clinical adoption [1–39, 41–44, 46–49, 51–64, 66–72, 75–81, 83–88, 91–94, 96, 97, 99, 100, 102].

Additional barriers include:

- **Lack of universal data standards:** Inconsistent structure, labeling, and exchange protocols impede interoperability and multi-center collaborations [17, 68, 77].
- **Global scalability and heterogeneity:** State-of-the-art AI and signal processing models often demonstrate reduced generalizability outside well-annotated, homogeneous cohorts, calling for improved domain adaptation, quality-aware fusion, and dynamic modality selection [32, 57, 71, 97].

- **Explainability:** Progress toward actionable and domain-relevant explanations is critical for supporting clinical accountability, user trust, and regulatory acceptance [6, 20, 75, 79].
- **Digital equity and privacy/security:** Variability in access to analytics, infrastructure, and digital literacy—both inter- and intra-regionally—risks entrenching health disparities. Concurrently, the transition to granular, continuous, interoperable data streams heightens privacy and security risks, necessitating progress in cryptography, differential privacy, and auditable governance [39, 68, 79, 81, 81, 86].
- **Responsible AI and collaborative research:** Ensuring responsible, bias-mitigated AI deployment requires regulatory harmonization and inclusive frameworks that bring together multidisciplinary stakeholders across research, clinical, technical, regulatory, and patient spheres [1–39, 41–44, 46–49, 51–64, 66–72, 75–81, 83–88, 91–94, 96, 97, 99, 100, 102].

To provide a clear overview, the following table summarizes the persistent challenges along with their identified impacts and key needs:

In summary, the synthesis of progress across biomedical signal processing, multimodal and cross-modal analytics, deep learning, IoT/real-time systems, synthetic data methodologies, and explainable AI marks a decisive step toward more precise, predictive, and participatory healthcare. Nevertheless, realization of this promise will depend upon concerted and sustained efforts to standardize, validate, explain, secure, and equitably distribute these innovations at scale—always guided by ethical, social, and clinical imperatives.

## References

- [1] O. Aerts, S. W. Griffiths, M. W. A. van Timmeren, and et al. 2020. The Image Biomarker Standardisation Initiative: Standardized Quantitative Radiomics for High Throughput Image-based Phenotyping. *Radiology* 295, 2 (2020), 328–338. doi:10.1148/radiol.2020191145
- [2] Ijaz Ahmad, Zhenzhen Liu, Lin Li, Inam Ullah, Sunday Timothy Aboyeji, Xin Wang, Oluwarotimi W. Samuel, Guanglin Li, Yuan Tao, Yan Chen, and Shixiong Chen. 2024. Robust Epileptic Seizure Detection Based on Biomedical Signals Using an Advanced Multi-View Deep Feature Learning Approach. *IEEE Journal of Biomedical and Health Informatics* 28, 10 (2024), To appear. doi:10.1109/JBHI.2024.3350919
- [3] R. Ahmadirohani, M. H. Karimpour, B. Rahimi, A. Malekzadeh-Shafaroudi, A. Beiranvand Pour, and B. Pradhan. 2018. Integration of SPOT-5 and ASTER satellite data for structural tracing and hydrothermal alteration mineral mapping: implications for Cu–Au prospecting. *International Journal of Image and Data Fusion* 9, 3 (2018), 239–253. doi:10.1080/19479832.2018.1469548
- [4] C. M. Alonso, M. Llop, C. Sargas, L. Pedrola, J. Panadero, D. Hervás, J. Cervera, E. Such, M. Ibáñez, R. Ayala, J. Martínez-López, E. Onecha, I. de Juan, S. Palanca, D. Martínez-Cuadrón, R. Rodríguez-Veiga, B. Boluda, P. Montesinos, G. Sanz, M. A. Sanz, and E. Barragán. 2019. Clinical Utility of a Next-Generation Sequencing Panel for Acute Myeloid Leukemia Diagnostics. *The Journal of Molecular Diagnostics* 21, 2 (2019), 228–240. doi:10.1016/j.jmoldx.2018.09.009
- [5] M. Antillon, M. McFarland, J. Ndung'u, E. Sutherland, M.E. Chappuis, S. Snijders, A. Hope, W.J. de Vlas, and E.M. Castaño. 2020. Cost-effectiveness of sleeping sickness elimination campaigns in the Democratic Republic of Congo: a transmission model-based economic evaluation. *medRxiv* (2020). doi:10.1101/2020.08.25.20181982
- [6] R. J. Appleton, B. C. Barnes, and A. Strachan. 2025. Data Fusion of Deep Learned Molecular Embeddings for Property Prediction. *arXiv preprint arXiv:2504.07297* (2025). <https://arxiv.org/abs/2504.07297>
- [7] F. Atilla, B. Klomberg, B. Cardoso, and N. Cohn. 2023. Background check: cross-cultural differences in the spatial context of comic scenes. *Multimodal Communication* 12, 3 (2023), 179–189. <https://www.degruyter.com/document/doi/10.1515/mc-2023-2010/html>
- [8] Mohammed Badawy, Nagy Ramadan, and Hesham Ahmed Hefny. 2023. Healthcare predictive analytics using machine learning and deep learning techniques: a survey. *Journal of Electrical Systems and Information Technology* 10 (2023). doi:10.1186/s43067-023-00108-y
- [9] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. Shinohara, et al. 2019. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. *Medical Image Analysis* 55 (2019), 80–91. doi:10.1016/j.media.2019.04.015
- [10] W. Barker, W. Chang, J. Everson, M. Gabriel, V. Patel, C. Richwine, and C. Strawley. 2024. The Evolution of Health Information Technology for Enhanced Patient-Centric Care in the United States: Data-Driven Descriptive Study. *Journal of Medical Internet Research* 26, 1 (2024), e59791. doi:10.2196/59791
- [11] J. A. Bateman and F. Schmidt-Borchering. 2018. The Communicative Effectiveness of Education Videos: Towards an Empirically-Motivated Multimodal Account. *Multimodal Technologies and Interaction* 2, 3 (2018), 59. <https://www.mdpi.com/2414-4088/2/3/59>
- [12] J. A. Bateman and C.-I. Tseng. 2023. Multimodal discourse analysis as a method for revealing narrative strategies in news videos. *Multimodal Communication* 12, 3 (2023), 261–285. doi:10.1515/mc-2023-2017
- [13] D. Benielli, H. Berry, and J. Morlier. 2022. scikit-multimodalllearn: A Python Toolbox for Multimodal Machine Learning. *Journal of Machine Learning Research* 23, 279 (2022), 1–6. <https://www.jmlr.org/papers/volume23/21-0791/21-0791.pdf>
- [14] Béatrice Biancardi, Maurizio Mancini, Brian Ravenet, and Giovanna Varni. 2024. Modelling the 'transactive memory system' in multimodal multiparty interactions. *Journal on Multimodal User Interfaces* 18, 1 (2024), 103–117. doi:10.1007/s12193-023-00399-w
- [15] F. Byrne. 2024. A diffractive approach to multimodal transcription: Materialising entanglements between humans and non-humans. *Multimodality & Society* (2024). doi:10.1177/26349795241291369
- [16] N. Bürgisser, E. Chalot, S. Mehouchi, C. P. Buclin, K. Lauper, D. S. Courvoisier, and D. Mongin. 2024. Large language models for accurate disease detection in electronic health records. *medRxiv* (2024). <https://www.medrxiv.org/content/10.1101/2024.07.27.24311106v1>
- [17] Kavyashree C., H. S. Vimala, and Shreyas J. 2024. A systematic review of artificial intelligence techniques for oral cancer detection. *Healthcare Analytics* 7 (2024), 100304. doi:10.1016/j.health.2024.100304
- [18] Christos Chatzichristos, Lauren Swinnen, Jaiver Macea, Miguel Bhagubai, Wim Van Paesschen, and Maarten De Vos. 2022. Multimodal detection of typical absence seizures in home environment with wearable electrodes. *Frontiers in Signal Processing* 2 (2022), 1014700. doi:10.3389/frsip.2022.1014700
- [19] Z. Chen, Y. Zheng, and J. C. Gee. 2024. TransMatch: A Transformer-Based Multilevel Dual-Stream Feature Matching Network for Unsupervised Deformable Image Registration. *IEEE Transactions on Medical Imaging* 43, 1 (2024), 15–27. doi:10.1109/TMI.2023.3324623
- [20] G. Dharmarathne, T. N. Jayasinghe, M. Bogawaththa, D. Meddage, and U. Rathnayake. 2024. A novel machine learning approach for diagnosing diabetes with a self-explainable interface. *Healthcare Analytics* 7 (2024), 100301. doi:10.1016/j.health.2024.100301
- [21] Hamdi Dibeklioglu, Elif Surer, Albert Ali Salah, and Thierry Dutoit. 2021. Behavior and usability analysis for multimodal user interfaces. *Journal on Multimodal User Interfaces* 15 (2021), 335–336. doi:10.1007/s12193-021-00372-0
- [22] A. Diez-Oliván, J. Del Ser, D. Galar, and B. Sierra. 2019. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. *Information Fusion* 50 (Jan. 2019), 92–111. doi:10.1016/j.inffus.2018.10.005
- [23] D. Ding, S. Li, B. Narasimhan, and R. Tibshirani. 2022. Cooperative learning for multiview analysis. *Nature Machine Intelligence* 4 (2022), 263–271. <https://www.nature.com/articles/s42256-022-00497-6>
- [24] I. J. Borges do Nascimento, M. S. Marcolino, H. M. Abdulazeem, I. Weerasekara, N. Azzopardi-Muscat, M. A. Gonçalves, and D. Novillo-Ortiz. 2021. Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies. *Journal of Medical Internet Research* 23, 4 (2021), e27275. doi:10.2196/27275
- [25] P. Dutta. 2018. An uncertainty measure and fusion rule for conflict evidences of big data via Dempster–Shafer theory. *International Journal of Image and Data Fusion* 9, 2 (2018), 187–201. <https://www.tandfonline.com/doi/full/10.1080/19479832.2017.1391336>
- [26] W. Feng, Q. Li, C. Wang, and J. s. Fan. 2025. Data Fusion for Full-Range Response Reconstruction via Diffusion Models. *arXiv preprint arXiv:2502.00795* (2025). <https://arxiv.org/abs/2502.00795>
- [27] S. Flynn, A. Haenel, F. Coughlan, S. Crilly, J. A. Leipsic, and J. D. Dodd. 2024. Cardiac CT, MRI, and PET in 2023: Exploration of Key Articles across Imaging and Multidisciplinary Journals. *Radiology* 313, 3 (2024), e240975. doi:10.1148/radiol.240975
- [28] F. Fraioli, N. Albert, R. Boellaard, I. Boscolo Galazzo, M. Brendel, I. Buvat, M. Castellaro, D. Cecchin, P. Aguiar Fernandez, E. Guedj, A. Hammers, Z. Kaplar, S. Morbelli, L. Papp, K. Shi, N. Tolboom, T. Traub-Weidinger, A. Verger, D. Van Weehaeghe, I. Yakushev, and H. Barthel. 2024. Perspectives of the European Association of Nuclear Medicine on the role of artificial intelligence (AI) in

Table 9: Persistent Challenges and Required Interventions for Scaling Biomedical AI in Healthcare

Challenge	Impact	Key Needs and Solutions
Standardization	Limits reproducibility, interoperability, clinical trust	Universal data formats, reporting standards, feature harmonization (e.g., Image Biomarker Standardization Initiative)
Global Scalability and Heterogeneity	Reduces accuracy in diverse/underrepresented groups	Domain adaptation, quality-aware fusion, data augmentation for imbalanced datasets
Explainability	Slows clinical adoption, impedes trust and accountability	Domain-relevant, actionable explanations (e.g., SHAP, LIME); regulatory guidance on XAI
Digital Equity and Privacy/Security	Exacerbates health disparities, threatens confidentiality	Policy-driven infrastructure investments, digital literacy programs, advanced privacy and security technologies
Responsible AI and Collaborative Research	Risks bias, misdeployment, ethical lapses	Inclusive stakeholder collaboration, rigorous validation, transparent and ethical governance

molecular brain imaging. *European Journal of Nuclear Medicine and Molecular Imaging* 51 (2024), 1007–1011. doi:10.1007/s00259-023-06553-1

[29] M. Giuffrè et al. 2023. Harnessing the power of synthetic data in healthcare. *npj Digital Medicine* 6 (2023), Article No. 198. doi:10.1038/s41746-023-00927-3

[30] P. Goyal, S. Sahu, S. Ghosh, and C. Lee. 2020. Cross-modal Learning for Multimodal Video Categorization. *arXiv preprint arXiv:2003.03501* (2020). <https://arxiv.org/abs/2003.03501>

[31] M. Habib, Z. Wang, S. Qiu, H. Liu, J. Qi, P. Kumar, and W. Chen. 2022. Machine Learning Based Healthcare System for Investigating the Association Between Depression and Quality of Life. *IEEE Journal of Biomedical and Health Informatics* 26, 6 (2022), 2795–2805. doi:10.1109/JBHI.2022.3152053

[32] Y. He, L. Zheng, X. Shi, and T. Huang. 2024. Efficient Modality Selection in Multimodal Learning. *Journal of Machine Learning Research* 25, 61 (2024), 1–53. <https://www.jmlr.org/papers/volume25/23-0439/23-0439.pdf>

[33] Y. Himeur, A. Ghanem, F. Alsalemi, R. Alhameed, F. Bensaali, and A. Amira. 2020. Data fusion strategies for energy efficiency in buildings: Overview, taxonomy, and future research directions. *Information Fusion* 64 (Oct. 2020), 99–120. <https://linkinghub.elsevier.com/retrieve/pii/S1566253520303158>

[34] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller. 2022. Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence. *Information Fusion* 79 (Nov. 2022), 263–278. <https://linkinghub.elsevier.com/retrieve/pii/S1566253521002050>

[35] A. Hope, C. F. Dietrich, L. Q. Zhou, D. L. McFarlane, R. J. Gillies, and H. Hricak. 2016. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* 278, 2 (2016), 563–577. <https://pubs.rsna.org/doi/10.1148/radiol.2015151169>

[36] P. Hu, Z. Huang, D. Peng, X. Wang, and X. Peng. 2023. Cross-modal retrieval with partially mismatched pairs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 8 (2023), 9595–9610. <https://ieeexplore.ieee.org/document/10050111/>

[37] Ali K. Ibrahim, Hanqi Zhuang, Emmanuelle Tognoli, Ali Muhamed Ali, and Nurgun Erdol. 2023. Epileptic seizure prediction based on multiresolution convolutional neural networks. *Frontiers in Signal Processing* 2 (2023), 1175305. doi:10.3389/frsip.2023.1175305

[38] S. Iino, R. Ito, K. Doi, T. Imaizumi, S. Hikosaka, and T. Y. Saito. 2018. CNN-based generation of high-accuracy urban distribution maps utilising SAR satellite imagery for short-term change monitoring. *International Journal of Image and Data Fusion* 9, 4 (2018), 298–312. doi:10.1080/19479832.2018.1491897

[39] P. Ivaturi, D. Chen, T. M. Talbert, Z. Jiang, W. Xu, G. Quer, M. Gadaleta, D. Jha, J. C. Lin, S. Santini, and P. Bonato. 2021. A Comprehensive Explanation Framework for Biomedical Time Series Classification. *IEEE Journal of Biomedical and Health Informatics* 25, 7 (2021), 2398–2408. doi:10.1109/JBHI.2021.3060997

[40] B. R. Jackson, C. Tarr, E. Strain, K. A. Jackson, A. Conrad, H. Carleton, L. S. Katz, S. Stroika, L. H. Gould, R. K. Mody, B. J. Silk, J. Beal, Y. Chen, R. Timme, M. Doyle, A. Fields, M. Wise, G. Tillman, S. Defibaugh-Chavez, Z. Kucerova, A. Sabol, K. Roache, E. Trees, M. Simmons, J. Wasilenko, K. Kubota, H. Pouseele, W. Klimke, J. Besser, E. Brown, M. Allard, and P. Gerner-Smidt. 2016. Implementation of Nationwide Real-time Whole-genome Sequencing to Enhance Listeriosis Outbreak Detection and Investigation. *Clinical Infectious Diseases* 63, 3 (2016), 380–386. doi:10.1093/cid/ciw242

[41] M. Jaritz, T.-H. Vu, R. De Charette, É. Wirbel, and P. Pérez. 2022. Cross-modal learning for domain adaptation in 3D semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 2 (2022), 1533–1544. <https://ieeexplore.ieee.org/document/9737217/>

[42] L. J. Jennings, M. E. Arcila, C. Corless, S. Kamel-Reid, I. M. Lubin, J. Pfeifer, R. L. Temple-Smolkin, K. V. Voelkerding, and M. N. Nikiforova. 2017. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *The Journal of Molecular Diagnostics* 19, 3 (2017), 341–365. doi:10.1016/j.jmoldx.2017.01.011

[43] S. Kjær Jensen and B. Schirmacher. 2024. Stronger together: Moving towards a combined multimodal and intermedial model. *Multimodality & Society* (2024). doi:10.1177/26349795241259606

[44] Yushan Jiang, Kanghui Ning, Zijie Pan, Xuyang Shen, Jingchao Ni, Wenchao Yu, Anderson Schneider, Haifeng Chen, Yuriy Nevmyvaka, and Dongjin Song. 2025. Multi-modal Time Series Analysis: A Tutorial and Survey. *arXiv preprint arXiv:2503.13709* (2025). <https://arxiv.org/abs/2503.13709>

[45] L. M. R. Kasbergen, E. de Bruin, F. Chandler, A. J. Menting, P. A. M. Kouwijzer, L. Schwanitz, V. Klaus, K. H. Adan, S. Markota, R. G. P. Alblas, et al. 2024. Multi-antigen serology and a diagnostic algorithm for the detection of arbovirus infections as novel tools for arbovirus preparedness in southeast Europe: a prospective observational study. *The Lancet Infectious Diseases* 24, 6 (2024), 689–698. doi:10.1016/S1473-3099(24)00654-6

[46] M. V. Kiang, E. T. Chin, B. Q. Huynh, L. A. C. Chapman, I. Rodríguez-Barraquer, B. Greenhouse, G. W. Rutherford, K. Bibbins-Domingo, D. Havlir, S. Basu, and N. C. Lo. 2021. Routine asymptomatic testing strategies for airline travel during the COVID-19 pandemic: a simulation study. *The Lancet Infectious Diseases* 21, 7 (2021), 929–938. doi:10.1016/S1473-3099(21)00134-1

[47] E. M. Kudalkar, N. A. M. Almontashiri, C. Huang, B. Anekella, M. Bowser, E. Hynes, R. Garlick, and B. H. Funke. 2016. Multiplexed Reference Materials as Controls for Diagnostic Next-Generation Sequencing: A Pilot Investigating Applications for Hypertrophic Cardiomyopathy. *The Journal of Molecular Diagnostics* 18, 6 (2016), 882–889. doi:10.1016/j.jmoldx.2016.07.005

[48] A. Kumar, P. Nandakishore, A. J. Gordon, E. Baum, J. Madhok, Y. Duanmu, and J. Kugler. 2025. Creation of an Open-Access Lung Ultrasound Image Database For Deep Learning and Neural Network Applications. *medRxiv* (2025). doi:10.1101/2025.05.09.25327337

[49] P. M. Kumar, U. P. Rao, P. Vijayakumar, N. Chilamkurti, D. Q. Phung, and N. Kumaravel. 2022. Clouds Proportionate Medical Data Stream Analytics for Internet of Things-Based Healthcare Systems. *IEEE Journal of Biomedical and Health Informatics* 26, 3 (2022), 973–982. <https://ieeexplore.ieee.org/document/9520255/>

[50] J. Lamb. 2018. To Boldly Go: Feedback as Digital, Multimodal Dialogue. *Multimodal Technologies and Interaction* 2, 3 (2018), 49. <https://www.mdpi.com/2414-4088/2/3/49>

[51] C. Lauri, A. Signore, A.W.J.M. Glaudemans, G. Treglia, O. Gheysens, R.H.J.A. Slart, R. Iezzi, N.H.J. Prakken, E.S. Debus, S. Honig, A. Lejay, and N. Chakfé. 2022. Evidence-based guideline of the European Association of Nuclear Medicine (EANM) on imaging infection in vascular grafts. *European Journal of Nuclear Medicine and Molecular Imaging* 49 (2022), 3430–3451. doi:10.1007/s00259-022-05769-x

[52] M. B. Lawson, W. Zhu, D. L. Miglioretti, T. Onega, L. M. Henderson, G. H. Rauscher, K. Kerlikowske, B. L. Sprague, E. J. A. Bowles, E. S. O’Meara, A. N. A. Tosteson, R. M. diFlorio Alexander, R. A. Hubbard, J. M. Lee, and C. I. Lee. 2025. Disparities in Standard-of-Care, Advanced, and Same-Day Diagnostic Services among Patients with Abnormal Screening Mammography. *Radiology* 314, 2 (2025), e241673. doi:10.1148/radiol.241673

[53] H. Laçi, K. Sevrani, and S. Iqbal. 2025. Deep learning approaches for classification tasks in medical X-ray, MRI, and ultrasound images: a scoping review. *BMC Medical Imaging* 25, 1 (2025), 156. <https://bmcmimedimaging.biomedcentral.com/articles/10.1186/s12880-024-01122-0>

[54] Hyeonhoon Lee, Hyun-Lim Yang, Hyuk Goo Ryu, Chiwon Jung, Yong Joon Cho, Seungbeom Yoon, Hyun-Kyung Yoon, and Hee-Chang Lee. 2023. Real-time machine learning model to predict in-hospital cardiac arrest using heart rate variability in ICU. *npj Digital Medicine* 6 (2023), Article no. 215. doi:10.1038/s41746-023-00960-2

[55] J. Li, P. Zhang, T. Wang, L. Zhu, R. Liu, X. Yang, K. Wang, D. Shen, and B. Sheng. 2024. DSMT-Net: Dual Self-Supervised Multi-Operator Transformation for Multi-Source Endoscopic Ultrasound Diagnosis. *IEEE Transactions on Medical Imaging* 43, 1 (2024), 64–75. doi:10.1109/TMI.2023.3324373

[56] X. Li, Y. Guo, S. Huang, F. Wang, C. Dai, J. Zhou, and D. Lin. 2025. A CT-based intratumoral and peritumoral radiomics nomogram for postoperative recurrence risk stratification in localized clear cell renal cell carcinoma. *BMC Medical Imaging* 25, 1 (2025), 167. doi:10.1186/s12880-024-01133-x

[57] Xiaoxiao Li, Liye Mei, Mengping Long, Jin Huang, Wei Yang, Xin Hao, Yi qiang Liu, Jinxuan Hou, Yu Xu, Fuling Zhou, Dupei Wang, Jianghua Wu, Taobao Hu, and Cheng Lei. 2025. Accurately matching serial whole slide images for tumor heterogeneity analysis. *Biomedical Signal Processing and Control* 87 (2025), 106825. doi:10.1016/j.bspc.2024.106825

[58] Z. Lin, S. Yu, Z. Kuang, D. Pathak, and D. Ramanan. 2023. Multimodality Helps Unimodality: Cross-Modal Few-Shot Learning with Multimodal Models. *arXiv preprint arXiv:2301.06267* (2023). <https://arxiv.org/abs/2301.06267>

[59] F. H. Linden, M. Boesen, E. Høegh-schmidt, J. U. Nybing, R. B. Mainz Hansen, L. H. Persson Kopp, M. N. Thomsen, and M. W. Brejnebol. 2025. External

- Validation of an Automated Segmentation Tool for Abdominal Fat Tissue using DIXON MRI: Data from the CutDM trial. *medRxiv* (2025). doi:10.1101/2025.05.14.25327583
- [60] D. Lindenberg. 2023. Multimodal meaning-making in student presentations: the impact of explicit feedback in a German as a foreign language classroom. *Multimodal Communication* 12, 3 (2023), 191–206. doi:10.1515/mc-2023-2011
- [61] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez. 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis* 42 (2017), 60–88. doi:10.1016/j.media.2017.07.005
- [62] Y. Liu, G. Li, and L. Lin. 2023. Cross-modal causal relational reasoning for event-level visual question answering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 10 (2023), 11624–11641. <https://ieeexplore.ieee.org/document/10146482/>
- [63] F. Ma, C. Liu, C. Aggarwal, J. Gao, and J. Han. 2020. Self-paced Multi-view Co-training. *Journal of Machine Learning Research* 21, 182 (2020), 1–31. <https://www.jmlr.org/papers/volume21/18-794/18-794.pdf>
- [64] S. J. Mataraso, C. A. Espinosa, D. Seong, S. M. Reincke, E. Berson, J. D. Reiss, Y. Kim, M. Ghanem, C.-H. Shu, T. James, Y. Tan, S. Shome, I. A. Stelzer, D. Feysaerts, R. J. Wong, G. M. Shaw, M. S. Angst, B. Gaudilliere, D. K. Stevenson, and N. Aghaepour. 2025. A machine learning approach to leveraging electronic health records for enhanced omics analysis. *Nature Machine Intelligence* 7 (2025), 293–306. doi:10.1038/s42256-024-00974-9
- [65] A. Matov. 2024. Urinary Biomarkers for Disease Detection. *medRxiv* (2024). <https://www.medrxiv.org/content/10.1101/2024.07.30.2431186v1>
- [66] Candy Olivia Mawalin, Shogo Okada, Yukiko I. Nakano, and Masashi Unoki. 2023. Personality trait estimation in group discussions using multimodal analysis and speaker embedding. *Journal on Multimodal User Interfaces* 17 (2023), 47–63. doi:10.1007/s12193-023-00401-0
- [67] Camille Mazzara, Pietro Sorrentino, Irene Altuna, Roma Maghsoudlou, Fabio Ferrarelli, Diego Sona, Luca Cocchi, Alessandro Gozzi, Carlo Miniussi, Silvia De Santis, Francesca Pistoia, Kenji Kansaku, Pau Erola, Gregoire Danaila, Carlo Cavedini, Barry A. Giesbrecht, Gordon Sajda, Rodolphe Jenatton, and Emmanuel Daucé. 2024. Mapping brain lesions to conduction delays: the next step for personalized brain models in Multiple Sclerosis. *medRxiv* (2024). doi:10.1101/2024.07.10.24310150 Preprint.
- [68] N. Mehta, Z. Pandit, and S. Shukla. 2019. Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study. *Journal of Biomedical Informatics* 100 (2019), Article 103311. doi:10.1016/J.JBI.2019.103311
- [69] F. Meng, T. Zhang, Y. Pan, X. Kan, Y. Xia, M. Xu, J. Cai, F. Liu, and Y. Ge. 2025. A deep learning algorithm for automated adrenal gland segmentation on non-contrast CT images. *BMC Medical Imaging* 25, 1 (2025), 142. doi:10.1186/s12880-025-01682-5
- [70] S. Mohammed, T.-H. Kim, R.-S. Chang, and C. Ramos. 2022. Guest Editorial: Data Analytics for Public Health Care. *IEEE Journal of Biomedical and Health Informatics* 26, 4 (2022), 1409–1410. <https://ieeexplore.ieee.org/document/9757035/>
- [71] T. Nguyen-Ky, H. D. Tuan, A. V. Savkin, M. N. Do, and N. T. T. Van. 2021. Real-Time EEG Signal Classification for Monitoring and Predicting the Transition Between Different Anaesthetic States. *IEEE Transactions on Biomedical Engineering* 68, 5 (2021), 1450–1458. doi:10.1109/TBME.2021.3053019
- [72] M. Nikolova-Simons, R. Keldermann, Y. Peters, W. Compagner, L. Montenij, Y. de Jong, and R. A. Bouwman. 2023. Predictive analytics for cardio-thoracic surgery duration as a stepstone towards data-driven capacity management. *npj Digital Medicine* 6 (2023). doi:10.1038/s41746-023-00938-0
- [73] I. N. Okeke and C. Ihekweazu. 2021. The importance of molecular diagnostics for infectious diseases in low-resource settings. *Nature Reviews Microbiology* 19 (2021), 547–548. doi:10.1038/s41579-021-00598-5
- [74] A. Pierangeli, O. Turriziani, M. Fracella, R. Campagna, F. Frasca, A. D'Auria, C. Scagnolari, P. Roberto, L. Cintì, G. D'Ettore, G. Ceccarelli, L. Petrarca, R. Nenna, F. Midulla, G. Galardo, and G. Antonelli. 2025. The added value of diagnostics to characterize age-specific patterns of respiratory viral infections and coinfections and to detect emerging threats. *BMC Infectious Diseases* 25 (2025), 404. doi:10.1186/s12879-025-10693-0
- [75] Mirco Planamente, Chiara Plizzari, Simone Alberto Peirone, Barbara Caputo, and Andrea Bottino. 2024. Relative Norm Alignment for Tackling Domain Shift in Deep Multi-modal Classification. *International Journal of Computer Vision* 132 (2024), 2618–2638. doi:10.1007/s11263-024-01998-9
- [76] E. H. P. Pooch, G. Agrotis, L. Cai, M. Emberton, T. T. Shah, H. U. Ahmed, R. G. H. Beets-Tan, S. Benson, T. Janssen, and I. G. Schoots. 2025. Semi-Supervised Learning in Prostate MRI Tumor Segmentation Approaches Fully-Supervised Performance on External Validation. *medRxiv* (2025). doi:10.1101/2025.05.13.25327456
- [77] R. A. Rasul, P. K. Saha, D. Bala, S. Karim, Md. I. Abdullah, and B. P. Saha. 2024. An evaluation of machine learning approaches for early diagnosis of autism spectrum disorder. *Healthcare Analytics* 7 (2024), 100293. doi:10.1016/j.health.2024.100293
- [78] M. P. Recht, L. M. White, J. Fritz, and D. L. Resnick. 2023. Advances in Musculoskeletal Imaging: Recent Developments and Predictions for the Future. *Radiology* 308, 2 (2023), e230615. doi:10.1148/radiol.230615
- [79] H. R. Roth, X. Zhuang, D. Shen, and C. Shen. 2022. Federated learning for medical image analysis. *Medical Image Analysis* 75 (2022), 102298. doi:10.1016/j.media.2021.102298
- [80] O. W. Samuel, Y. Z. Kulwa, G. Li, R. N. Khushaba, and S. S. Ge. 2022. Multiresolution Dual-Polynomial Decomposition Approach for Optimized Characterization of Motor Intent in Myoelectric Control Systems. *IEEE Transactions on Biomedical Engineering* 69, 10 (2022), 3122–3133. doi:10.1109/tbme.2022.3221799
- [81] T. Schäck, M. Muma, M. Feng, C. Guan, and A. M. Zoubir. 2018. Robust Nonlinear Causality Analysis of Nonstationary Multivariate Physiological Time Series. *IEEE Transactions on Biomedical Engineering* 65, 6 (2018), 1213–1225. <https://ieeexplore.ieee.org/document/7934390>
- [82] Gizachew Mulu Setegn and Belayneh Endalamaw Dejene. 2025. Explainable AI for Symptom-Based Detection of Monkeypox: a machine learning approach. *BMC Infectious Diseases* 25 (2025). doi:10.1186/s12879-025-10738-4
- [83] Chengdong Shi, Kai Yu, Yu Hu, Yuantao Wang, Fan Bu, Ji Lu, and Weigang Wang. 2024. Diagnostic Efficacy of Various Imaging Modalities Across Different Stages of Prostate Cancer: A Network Meta-Analysis of Diagnostic Studies. *medRxiv* (2024). doi:10.1101/2024.09.28.24314285
- [84] F. Sollini, R. Antunovic, A. Chiti, and H. J. W. L. Aerts. 2021. PET Radiomics in NSCLC: State of the Art and a Proposal for Harmonization of Methodology. *Radiology* 298, 2 (2021), 548–560. doi:10.1148/radiol.2021200616
- [85] A. Song, L. Xu, L. Wang, B. Wang, X. Yang, B. Xu, B. Yang, and S. E. Greenwald. 2022. Automatic coronary artery segmentation of CCTA images with an efficient feature-fusion-and-rectification 3D-UNet. *IEEE Journal of Biomedical and Health Informatics* 26, 8 (2022), 4044–4055. <https://ieeexplore.ieee.org/document/9405480>
- [86] Min-Seo Song and Seung-Bo Lee. 2024. Comparative study of time-frequency transformation methods for ECG signal classification. *Frontiers in Signal Processing* 4 (2024). doi:10.3389/frsip.2024.1322334
- [87] S. S. Srinivas and V. Runkana. 2024. Cross-Modal Learning for Chemistry Property Prediction: Large Language Models Meet Graph Machine Learning. *arXiv preprint arXiv:2408.14964* (2024). <https://arxiv.org/abs/2408.14964>
- [88] S. R. Stahlschmidt, B. Ulfenborg, and J. Synnergren. 2023. Multimodal deep learning for biomedical data fusion: a review. *Nature Machine Intelligence* 5 (2023), 351–362. <https://www.nature.com/articles/s42256-023-00629-w>
- [89] A. J. Sundermann, J. Chen, P. Kumar, A. M. Ayres, S. T. Cho, C. Ezeonwuka, M. P. Griffith, J. K. Miller, M. M. Mustapha, A. W. Pasculle, M. I. Saul, K. A. Shutt, V. Srinivasa, K. Waggle, D. J. Snyder, V. S. Cooper, D. Van Tyne, G. M. Snyder, J. W. Marsh, A. Dubrawski, M. S. Roberts, and L. H. Harrison. 2022. Whole-Genome Sequencing Surveillance and Machine Learning of the Electronic Health Record for Enhanced Healthcare Outbreak Detection. *Clinical Infectious Diseases* 75, 3 (2022), 476–482. doi:10.1093/cid/ciab946
- [90] A. J. Sundermann, P. Kumar, M. P. Griffith, K. D. Waggle, V. Rangachar Srinivasa, N. Raabe, E. G. Mills, H. Coyle, D. Ereifej, H. M. Creager, A. Ayres, D. Van Tyne, L. L. Pless, G. M. Snyder, M. Roberts, and L. H. Harrison. 2025. Real-Time Genomic Surveillance for Enhanced Healthcare Outbreak Detection and Control: Clinical and Economic Impact. *Clinical Infectious Diseases* (2025). doi:10.1093/cid/ciaf216
- [91] D. van Kampen and J. Wildfeuer. 2024. Cohesion in physical movie releases: developing a framework for the multimodal analysis of complex three-dimensional packaging. *Multimodality & Society* (2024). doi:10.1177/26349795241306993
- [92] A. Vobecky, D. Hurych, O. Siméoni, S. Gidaris, A. Bursuc, P. Pérez, and J. Sivic. 2025. Unsupervised Semantic Segmentation of Urban Scenes via Cross-Modal Distillation. *International Journal of Computer Vision* 133, 6 (2025), 3519–3541. doi:10.1007/s11263-024-02320-3
- [93] Jiaqi Wang, Hanqi Jiang, Yiheng Liu, Chong Ma, Xu Zhang, Yi Pan, Mengyuan Liu, Peiran Gu, Sichen Xia, Wenjun Li, Yutong Zhang, Zihao Wu, Zhengliang Liu, Tianyang Zhong, Bao Ge, Tuo Zhang, Ning Qiang, Xintao Hu, Xi Jiang, Xin Zhang, Wei Zhang, Dinggang Shen, Tianming Liu, and Shu Zhang. 2024. A Comprehensive Review of Multimodal Large Language Models: Performance and Challenges Across Different Tasks. *arXiv preprint arXiv:2408.01319* (2024). <https://arxiv.org/abs/2408.01319>
- [94] M. Wang, S. Li, T. Zheng, N. Li, Q. Shi, X. Zhuo, R. Ding, and Y. Huang. 2022. Big Data Health Care Platform With Multisource Heterogeneous Data Integration and Massive High-Dimensional Data Governance for Large Hospitals: Design, Development, and Application. *JMIR Medical Informatics* 10, 4 (2022), e36481. doi:10.2196/36481
- [95] T. Weitzel, C. Pérez, D. Tapia, P. Legarraga, and L. Porte. 2021. SARS-CoV-2 rapid antigen detection tests. *The Lancet Infectious Diseases* 21, 8 (2021), 1067–1068. doi:10.1016/S1473-3099(21)00249-8
- [96] O. Westerland, A. Amlani, C. Kelly-Morland, M. Fraczek, K. Bailey, M. Gleeson, I. El-Najjar, M. Streetly, P. Bassett, G. J. R. Cook, V. Goh, London Myeloma Imaging Research Group at Guy's & St Thomas' Hospital, and King's College London. 2021. Comparison of the diagnostic performance and impact on management of 18F-FDG PET/CT and whole-body MRI in multiple myeloma. *European Journal*

- of Nuclear Medicine and Molecular Imaging* 48 (2021), 2558–2565. doi:10.1007/s00259-020-05182-2
- [97] Q. Xie, Y. Li, N. He, M. Ning, K. Ma, G. Wang, Y. Lian, and Y. Zheng. 2024. Unsupervised Domain Adaptation for Medical Image Segmentation by Disentanglement Learning and Self-Training. *IEEE Transactions on Medical Imaging* 43, 1 (2024), 4–14. doi:10.1109/TMI.2023.3324198
- [98] X. Yang and J. Li. 2025. Learn from the SARS-CoV-2 nucleic acid test to increase the experience of dealing with the 'disease X'. *BMC Infectious Diseases* 25 (2025). <https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-025-10991-7>
- [99] A. Dehghani Zahedani, T. McLaughlin, A. Veluvali, N. Aghaeepour, A. Hoseinian, S. Agarwal, J. Ruan, S. Tripathi, M. Woodward, N. Hashemi, and M. Snyder. 2023. Digital health application integrating wearable data and behavioral patterns improves metabolic health. *npj Digital Medicine* 6 (2023). doi:10.1038/s41746-023-00956-y
- [100] Q. Zhang, Y. Wei, Z. Han, H. Fu, X. Peng, C. Deng, Q. Hu, C. Xu, J. Wen, D. Hu, and C. Zhang. 2024. Multimodal Fusion on Low-quality Data: A Comprehensive Survey. *arXiv preprint arXiv:2404.18947* (Nov. 2024). <https://arxiv.org/abs/2404.18947>
- [101] Wenwu Zhu, Xin Wang, and Hongzhi Li. 2020. Multi-modal Deep Analysis for Multimedia. *IEEE Transactions on Circuits and Systems for Video Technology* (2020). doi:10.1109/TCSVT.2019.2940647
- [102] X.-F. Zhu, T. Xu, Z. t. Liu, Z. Tang, X.-J. Wu, and J. Kittler. 2025. UniMod1K: A Dataset for Multimodal Learning with Vision, Depth, and Language. *International Journal of Computer Vision* 133, 6 (2025). <https://link.springer.com/journal/11263/volumes-and-issues/133-6> Special Issue on Multimodal Learning.