Computing&AI Connect
ISSN: 3104-4719
2025, Vol. 2, Cite as: doi:10.x/journal.x.x.x

SCIFINITI PUBLISHING

OPEN ACCESS

Review Article

Computing and AI in Genomics-Driven Precision Oncology: Methods, Trials, and Translation

Arbiona Shukulli[1], shukulliarbiona@gmail.com, arbionashukulli@precacollege.org,

0009-0002-4354-3200

Rahaf M. Ahmad[2*], 700039113@uaeu.ac.ae,

0000-0002-7531-5264

1- Preca College, Rruga Pandeli Vangjeli 5, Korçë, Albania
2- Department of Genetics and Genomics, College of Medicine and Health Sciences, United Arab Emirates University, United Arab Emirates

Abstract

Precision Oncology has brought about a complete paradigm shift in cancer treatment approaches from Histology-guided regimes to Genomic-anchored, precision-cancer therapies. This paper reviews the major milestones in basic translational work in cancer genomics, evolution in clinical trial designs, and translational activities led by top Universities such as Harvard, MIT, Oxford, and Cambridge. This paper describes the major translational breakthroughs such as The Cancer Genome Atlas (TCGA) project, Functional Genomics with CRISPR, Live Sequencing platforms like MatchMiner, and underscores the need for their collective advancement in stratifying patients and personalizing treatments. This paper also showcases how cases from major institutions have aided in integrating multi-omics, adaptive clinical trials, and ethical AI approaches in the realm of research and clinical practice. Digital Twin models in MIT, GenOMICC in Oxford, and Spatial Omics in Cambridge demonstrate diversified, mutually supportive approaches in the realm of translational precision medicine. Emerging approaches such as Single-Cell Sequencing, Spatial Omics, AI-enriched clinical trials demonstrate an imminent future marked by learning health platforms. The paper also underscores the impending issues in Equity, Harmonization, and variant understanding, contrary to the basic scientific breakthroughs. Disparities in clinical trials and a lack of representation in diverse populations might accentuate global health inequities. We believe

Precision Oncology mustering pioneer breakthroughs in Technological, Governance, Political, and Ethical domains. This paper also offers an inclusive blueprint in demarcating the limitations in Precision Oncology, intersecting Genomic Understanding, innovative clinical trials, and Socio-politics, prying into pioneer breakthroughs in treatments from renowned institutions to universally practical approaches in global health ecosystems.

Keywords: precision oncology; genomic medicine, adaptive clinical trials, artificial intelligence in cancer, translational research.

Highlights:

- Genomic profiling has revolutionised cancer classification, risk stratification, and treatment selection.
- Adaptive trial designs (e.g., basket, umbrella, and platform trials) are key to matching therapies to tumour genomics.
- Harvard, MIT, Oxford, and Cambridge lead in translational precision oncology through innovation in diagnostics, AI, and ethical frameworks.
- Barriers such as variant interpretation, data equity, and global access remain critical challenges.
- The future of precision oncology lies in real-time multi-omics integration, federated data systems, and AI-driven trial governance.

1. Introduction:

1.1 The Rise of Genomic Precision in Oncology

In the last 20 years advances in molecular profiling, next-generation sequencing (NGS) and computational biology have dramatically changed the oncology landscape. As a result of precision oncology, which utilizes genomic, transcriptomic and epigenetic data to make decisions about therapies, the standards of care for many cancer types - including metastatic breast cancer, cholangiocarcinoma, neuroendocrine tumors, and gliomas - have been rewritten. Studies like The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) helped to facilitate this transition from tumor classification using histopathology, to tumor/molecular subtypes and actionable mutations [1], [2]. Today, large-scale profiling of tumor genomes is becoming a standard part of routine oncology practice, providing new opportunities for targeted therapy, and early diagnosis, and real-time monitoring by liquid biopsies [3], [4]. Yet, although we are ushering in an era of precision oncology, clinical translation of genomic knowledge to therapy is still faced with many challenges. There is the troubled history of targeted agents failing late in developmental trials, that are then hampered by acquired resistance. It is troubling to realize a lack of representation of minority population genomic data [5], [6]. Moreover, understanding tumor heterogeneity both between patients (inter-subject) and within a patient's tumor sample (intra-subject) is exemplified by the limitations of tumor tissue and single site biopsies and highlights a need for longitudinal and multi-modal data acquisition [7]. This review presents an in-depth discussion on the translational path of precision oncology, incorporating the advances in genomics and the movement of clinical trials. Through institutional insights culled from Harvard Medical School, Broad Institute, MIT, University of Oxford, and University of Cambridge, this discussion defines how such hotbeds of excellence have brought major breakthroughs in their respective genomics, in addition to establishing innovative approaches in the design of clinical trials. At Harvard and the Broad, programs like MSK-IMPACT and OncoPanel have allowed for large-scale integration of clinical sequencing, while also contributing to national scale initiatives, such as NCI-MATCH [8], [9]. MIT's interdisciplinary convergence model has resulted in, among other innovations, novel approaches to nanoparticle drug delivery and organoid-based pre-screening for clinical trials [10]. Oxford and Cambridge have been integral in pioneering

adaptive and biomarker-enhanced trial designs through programs like the UK 100,000 Genomes Project, FOCUS4, and TRACERx [11], [12]. We argue that precision oncology should no longer be regarded as a linear process from mutation to drug. Instead, we must view it as a dynamic, iterative process that takes advantage of real-time molecular information, AI-assisted process, and patient-centred endpoints. The viability of systems under this framework does not only depend on scientific advances, but also there is a point of need for infrastructure, regulatory, and equity considerations. There is a need to overcome genomic data interpretation bottlenecks, harmonise international trial methods, and broaden access to molecular diagnostics, especially in low- and middle-income contexts. By putting together a range of data across sub-disciplines and institutions, this review provides strategic and academic commentary on how genomics led precision oncology is being implemented through cancer clinical trials, in a historical analysis. Planning beyond what has been accomplished, we want to try to address what will threaten genomic insights to translate into impactful, measurable clinical outcome; especially in relation to pan-cancer, biomarker-driven, or AI-supported trials.

2. Genomic Foundations of Precision Oncology

Molecular characterisation of cancer has progressed from identifying recurrent mutations in protein coding genes to revealing the considerable complexity of the cancer genome and its functional consequences. Genomics has facilitated the dis-assembly of tumour biology with an unparalleled resolution, revealing layers of somatic mutations, copy number variation, structural rearrangements, chromatin organisation and epigenomic landscapes that interact to drive oncogenesis.

2.1 Comprehensive Cataloguing of Somatic Variants

The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) ushered in the age of large-scale and systematic characterisation of a variety of tumour types, resulting in massive public warehouses of genomic, transcriptomic and epigenomic data. Integrative analyses of over 33 cancer types from TCGA revealed recurrently mutated genes such as TP53, PIK3CA, and KRAS, novel fusion events, and widespread dysregulation of gene expression via epigenetic silencing through aberrant DNA methylation and enhancer hijacking [2, 13]. In concert, these studies provided a springboard for pan-cancer analyses, beginning to understand commonalities and differences among tumour types beyond that of tissue of origin.

The Cancer Cell Line Encyclopedia (CCLE) from the Broad Institute and the Genomics of Drug Sensitivity in Cancer (GDSC) resource from the Sanger Institute provide systems-level connectivity between genomic alterations and pharmacological response across hundreds of human cancer cell lines [14], [15]. The datasets have helped not only to guide drug repurposing efforts, but also to develop and train machine learning methods to predict therapeutic response based on mutational profiles. Utilizing the Mutational Signatures Framework developed by the Wellcome Centre for Human Genetics at Oxford, several mutagenic processes have been identified: UV light (Signature 7), activity from the APOBEC family of deaminases (Signatures 2 and 13), and defective DNA mismatch repair (Signature 6) are now included in clinical genomics workflows to provide mechanistic understanding and support treatment choices, including the use of immune checkpoint inhibitors in tumours deficient in mismatch repair [16]. Cambridge has developed computational methods that include algorithms for clonal deconvolution (e.g., PyClone, SciClone) through the CRUK Cambridge Institute and European Bioinformatics Institute (EBI). These algorithms have provided new insights concerning clonal architecture and evolutionary patterns as the patient is placed under treatment pressure [17]. Despite these advances, genomic datasets including TCGA carry a bias toward individuals of European ancestry and while these assertions have improved generalizability across populations, limited availability is still a significant barrier for equitable clinical application. The underrepresentation of genomic analyses from diverse human populations is a major issue with respect to global health equity. However, there have been several recent developments such as the Pan-Cancer Analysis of Whole Genomes (PCAWG) project which has included whole-genome analyses and identified mutations in the noncoding genome, complex structural variations, and regulatory alterations that were previously untapped with exome sequencing.

2.2 Functional Genomics: From Mutation to Mechanism

To move from mutation catalogues to mechanistic insight, high-throughput functional screens are essential. CRISPR-Cas9 knockout libraries have allowed for genome-wide studies of gene essentiality, identifying not only genetic dependencies but also synthetic lethal relationships (e.g., BRCA-deficient cells when perturbed with PARP inhibitors) or context-specific dependencies [18]. The Broad Institute developed the Dependency Map (DepMap) project to examine CRISPR and RNAi data alongside gene expression, mutation status and drug response to characterise lineage-specific vulnerabilities or pan-cancer genetic dependencies. This map

of cancer gene dependencies has already been used to design selective inhibitors against genes that are essential only in specific tumour types, such as WRN helicase in microsatellite instability-high (MSI-H) tumours and STAG2 in Ewing sarcoma [19]. At MIT, CRISPR screens have already been adapted to combine single-cell RNA sequencing readouts with pooled combinatorial perturbations in order to identify buffering networks and insights into gene-gene interactions that are critically important for therapy resistance [20]. Simultaneously, at Oxford's Target Discovery Institute and Cambridge's Gurdon Institute, the same groups have started first-of-their-kind functional screens in 3D organoid cultures and patient-derived xenograft (PDX) cultures to better mimic the heterogeneity of living tumours in vitro [21]. Furthermore, AI-enabled models have recently started to combine existing CRISPR screen outputs with chromatin architecture and transcriptomic landscapes to develop models of regulatory interactions and regulatory vulnerabilities in specific contexts (e.g., GraphReg, CrisprBrain). These types of models allow for prioritisation of not only noncoding regulatory elements but also of synthetic lethal pairs (including, in rare or low frequency contexts).

2.3 Clinical-Grade Genomic Diagnostics and Decision Support

Incorporating genomic information into clinical workflows must be done using high fidelity, regulatory-grade assays for genomic cancer diagnostics. For example, MSK-IMPACT and Harvard's OncoPanel are hybrid-capture NGS panels that can identify somatic mutations, CNVs and rearrangements in medically actionable genes. Results are enriched with trial eligibility, utilizing platforms such as MatchMiner, which connect with prospective patients to studies relevant to their circumstance [9]. The NHS genomics offer has been developed to encompass whole genome sequencing (WGS) as a diagnostic tool in its cancer services through the 100,000 Genomes Project at Oxford and Cambridge, and produce curated calls for somatic and germline mutations to support both treatment and familiy risk assessment. The datasets also enable longitudinal EHRs so that correlative outcome analyses are possible [22]. An area of significant bottleneck within clinical genomics is the level of forced classification of variants of uncertain significance (VUS); this generates a statistically significant proportion of the results reported to patients. Databases such as ClinVar, CIViC, and OncoKB exist to curate the pathogenicity and therapeutic actionability of medically salient results however the classification process remains subjective and ad hoc. Ensemble AI models like REVEL and PathoMAN are being trained on large scale annotations to help with automated classification

of VUS and to streamline classification based on functional, structural, and evolutionary indications [23]. VUSs remain a barrier to clinical adoption. Proteome-wide predictors such as AlphaMissense and PrimateAI-3D provide scalable computational scoring of missense variants [24], [25]. Experimental multiplexed assays-including saturation genome editing-can classify thousands of variants in parallel, providing empirical benchmarks [26]. A combined workflow where computational predictors triage variants and multiplex assays confirm high-priority genes can accelerate reclassification in BRCA1/2 and other actionable cancer predisposition genes. Interpretation bottlenecks also limit tumour board scalability. The Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT and Big Data Institute at Oxford are developing AI models to automate variant classification, prioritise targets, and predict response with multi-modal inputs. These products will be integrated into molecular tumour boards and the decision support interface.

2.4 Emerging Technologies: Single-Cell and Spatial Multi-Omics

Cancer is spatially and temporally heterogeneous. Single-cell RNA sequencing (scRNA-seq) allowed the identification of rare subpopulations (e.g., drug-tolerant persisters) and lineage trajectories as they transition through epithelial–mesenchymal transition (EMT) states here at the Broad Institute [27]. Spatial transcriptomics platforms like Slide-seq and 10x Genomics Visium are mapping the architectural relationships between cancer cells, stroma, and immune infiltrates. At Cambridge, spatial multi-omics applied in colorectal and glioblastoma have found immune exclusion zones and hypoxic niches that correlated with resistance [28]. The concepts of barcoded nanoparticles applied to in vivo multiplexed drug screening at MIT afford opportunities to perform functional phenotyping in the native microenvironment. These technologies will inform patient stratification and trials [29]. Recent developments in the field of artificial intelligence (AI) have included the scVI (single-cell variability inference) and DeepCell, which model integrated datasets that consist of spatial transcriptomics data, digital pathology, and single-cell epigenomics. These models can show how gene expression changes over time, how cells interact with each other, and how spatial architectures work. They are a new type of tool for dividing trials into groups based on the topologies and microenvironmental status of the tumors.

3. Translating Genomic Insight into Clinical Trial Design

Adding genomic data to the structure of cancer clinical trials is a huge step forward in the development of new drugs. Historically, oncology trials worked primarily off of a histology-based stratification framework, which required the tumour to be of a particular type and stage, and did not account for the underlying molecular heterogeneity that is responsible for therapeutic response. Precision oncology trials do employ genomic biomarkers-mutational signatures, gene fusions, transcriptomic profiles-to assign patients to the therapies that are most likely to be successful.

3.1 From Histology-Based to Biomarker-Driven Trials

Histology-agnostic trial designs harmonize with the molecular complexity of cancer with patents quite broadly based on actionable mutations without restrictions based on the tumour type (basket trials, such as NCI-MATCH and ASCO TAPUR trials, the Dana-Farber Cancer Institute (Harvard pioneered) basket trials combine patients diagnosed with multiple type of cancer in which there is a shared actionable mutation [8], [30]. Basket trials are meant to alleviated the limitations of single-histology, as oncogenic drivers can cross tissue origin which might allow for wider eligibility and faster accrual. Umbrella trials (in the same vein) like Lung-MAP assigned patients who had the same tissue diagnosis (and were often met with the same treatment approach, e.g. NSCLC) through different arms based their mutation profiles [31]. The nature of these trials can highlight one of certain heterogeneity through a single tumor type, while accommodating a platform approach for rapid evaluation of multiple targeted therapies within a shared enrollment infrastructure. Many trials have introduced adaptive trial features (e.g. Bayesian arm expansion) and pre-specified interim analysis which allows for early stopping for futility or success. Adaptive approaches bring an efficiency to resource utilisation. Oxford's FOCUS4 trial in colorectal cancer is one of the first adaptive trials with genomic stratification and treatment arms that could be opened, or closed, based on interim data [32]. In a similar way, Cambridge and UCL's TRACERx lung cancer study tracks clonal evolution over time while also incorporating longitudinal genomics to guide treatment adaptation [12]. These types of trials have made real-time sequencing, adaptive randomisation, and interim futility analysis pillars of contemporary, and modern trial designs. Additionally, hybrid trial designs that enable incorporation of additional layers of data such as epigenetic profiles, immune landscapes, spatial transcriptomics, are being employed to allow for multi-

modal stratification. DeepTrial (Stanford), for example, collects these data streams, together with AI, to dynamically recommend treatment arms and stratification logic, framed against real-time patient-specific features [33].

3.2 Real-Time Sequencing and Dynamic Eligibility

Quick turnaround in NGS profiling must be a prerequisite to eligibility of biomarkers. The MatchMiner platform built at DFCI integrates patient genomic information with publicly available clinical trial protocols to allow AI algorithms to identify potential matches during a real time search process [34]. This method has shown success in achieving higher enrolment efficiencies, and curtailing the lag time between molecular diagnosis and trial initiation. Genomic data pipelines from Genomics England connected to NHS records in Oxford and Cambridge will allow matched patients to be identified, as soon as actionable alterations are detected. Cohorts such as these benefit from being integrated using data standards like GA4GH (Global Alliance for Genomics and Health) so that trial sites can communicate with each other. Profiling is also being aided by the availability of liquid biopsy tests which allow for the detection of ctDNA mutations in the real-time context of patient treatment. Liquid biopsies facilitate on-treatment monitoring and ,in some instances, real-time treatment switching [35]. Clinical trials such as DYNAMIC (Designation of cANcer through Diagnostic Imaging) in colorectal cancer, and the B-F1RST trial in non-small cell lung cancer (NSCLC) are now utilising these methods of real-time stratification and treatment direction [36], [37]. For example, acquired EGFR T790M mutation patients could be switched to osimertinib arms mid-treatment when the mutation is diagnosed via ctDNA, allowing precision therapy to develop in parallel with tumour biology. This is an example of "real-time eligibility," whereby eligibility criteria are variable and governed by the molecular status of the tumour over time.

3.3 Overcoming Barriers: Equity, Interpretation, and Scalability

There remain challenges, however, despite relative technical successes. Limited access remains as over-representation of ethnic minority and rural populations limits generalisability [5]. This is exacerbated by issues regarding access to sequencing infrastructure and those institutions conducting the trial research These institutions, such as Harvard and MIT, are piloting mobile phlebotomy units and digital consent processes to decentralise the recruitment process and potentially eliminate geographic and socioeconomic impediments. The GenOMICC project in

Oxford also represents an example of strive to include different ancestries to include severe cancer phenotypes and give examples of inclusive trial design. Furthermore, linkage through with biobank initiatives (eg, UK Biobank and All of Us Research Program) provide an opportunity to perform retrospective genomic profiling that relates to outcomes data.

3.4 Future Directions: AI-Driven Master Protocols

The future of precision oncology trials lies within algorithmically-enhanced, tumor-agnostic, and continuously adaptive master protocols that can account for the complexities associated with cancer biology and the continued influx of new molecular information. Master protocols established through the Precision Cancer Consortium (Harvard-affiliated) and the WISDOM trial framework are being pilot-tested to enable modular arm reconfiguration, continuous enrollment, and adaptive cohorting based on newly available biomarker or response to treatment data. The DARPA-enabled Intelligent Trial Design program at MIT is constructing self-optimizing systems to aid in the allocation of patients across the arms, leveraging omics data, patient clinical trajectories, and patient-reported outcomes [37]. These frameworks involve reinforcement learning and Bayesian optimization and help provide knowledge from prior arms to inform the structure of subsequent trials in real time [38]. The contributions of Oxford and Cambridge to the PAN-COVID cancer study has also informed frameworks for real-world trial extensions, where we are able to make adaptive changes to standard-of-care pathways without compromising scientific rigor. However, simultaneously, scientists and organizations implement federated learning approaches to train AI models without transferring patient data between hospitals. This helps ensure patient data privacy and overcomes obstacles in establishing collaborations between institutions, which makes it difficult for them to work together [38]. These developments make the integration of adaptive trials in the hospital treatment paradigm possible, with faster validation and support for more patient numbers. Apart from enhancing success rates and timelines pertaining to the validation of adaptive trials, such trials also allow, along with other factors, every person, including patients, to benefit from trials with reduced variation in availability, making patient satisfaction one of the design elements. The convergence of insights from genes, AI, and adaptive trials is revolutionizing the practice of clinical trials, making way for the realization of the full potential of precision oncology.

4. AI, Data Integration, and the Future of Precision Trial Design

There have been major breakthroughs in the past few decades in the field of medicine, thanks to scientific-technology development, also known as precision medicine. This branch of medicine is also revolutionizing diagnostic, treatment, and patient tracking approaches by integrating with artificial intelligence (AI) in the diagnosis of chronic and complex diseases accurately [39]. One such disease is cancer. This disease is complex, diverse, and also very causal in terms of morbidity and mortality globally [40], [41]. Moreover, this disease also varies significantly from person to person in terms of type, stage, and response to treatment [42]. Since this diversity exists, treatment cannot be standardized. Due to such diversity, treatment of such diseases has been difficult even for doctors ever since [40]. However, AI integration in such diseases has eased quite a few difficulties significantly [43]. Within the past 32 years, the mortality rate from such diseases has decreased by 33%. Advanced science in such diseases, especially the evolution of precision medicine in such diseases, has been the cause of such improvement in such diseases. Advanced science in such diseases has, in fact, progressed in terms of more effective treatment and more specific patient treatment by integrating with AI. AI helps in detecting hidden patterns in images, estimating the possible advancement of diseases, proposing treatments, and identifying whether the patient is eligible enough to enroll in clinical trials or not [44]. Indeed, in 2021, 71 AI-enabled devices were approved by the FDA. Moreover, more than 80% of such devices were used in the diagnosis of cancers. These devices were mostly utilized in radiology, pathology, and radiation oncology in cancers, especially solid cancers such as cancers in the breasts, lungs, and prostate [42], [44]. Even though such devices have been significantly utilized in terms of accuracy improvement in decision-making, their complete usage due to irregularities in such devices, even in their usage, has been impeded due to possible bias in algorithms along with their limitations in terms of efficiency in estimations, along with higher burdens on such healthcare systems [44].

4.1 AI in Patient Stratification and Trial Matching

One of the most influential applications of AI in precision oncology is in stratifying patients eligible for clinical trials, based on their genomic, image, and clinical information. The complexity associated with cancer, both intra- and inter-tumoral, makes conventional patient stratification in clinical trials notoriously difficult, typically causing the exclusion of patients

with unique profiles. AI helps overcome this hurdle by identifying hidden patterns in high-dimensional data, making patient stratification more accurate and inclusive [45], [46], [47].

The DeepPatient model, with data from more than 700,000 electronic health records, demonstrated the ability of unsupervised learning to identify patient properties and disease patterns in patient cohorts [48]. DeepMatch at Memorial Sloan Kettering Cancer Center evaluates genetic and clinical characteristics in real-time to accurately match patients to ongoing studies. Furthermore, AI-driven instruments used in the I-PREDICT study produced "matching scores" derived from integrated multi-omics and biomarker data, facilitating treatment selection and associated with enhanced progression-free survival [49]. Radiomic features fall into four IBSI-standardized classes: (i) shape/morphology; (ii) first-order intensity; (iii) texture (GLCM, GLRLM, GLSZM, NGTDM); and (iv) filtered/wavelet features. To minimize compute and redundancy, features are filtered for reproducibility, highly correlated variables are removed (e.g., $|r|>0.9$), and embedded selection methods such as LASSO or mRMR are applied. Cross-validation determines the smallest performant set, enabling efficient and generalizable modelling [50], [51].

Figure 1 below illustrates the clinical challenge: the complexity of radiomic data increases with metastatic disease, requiring standardized AI-driven pipelines for effective trial stratification.

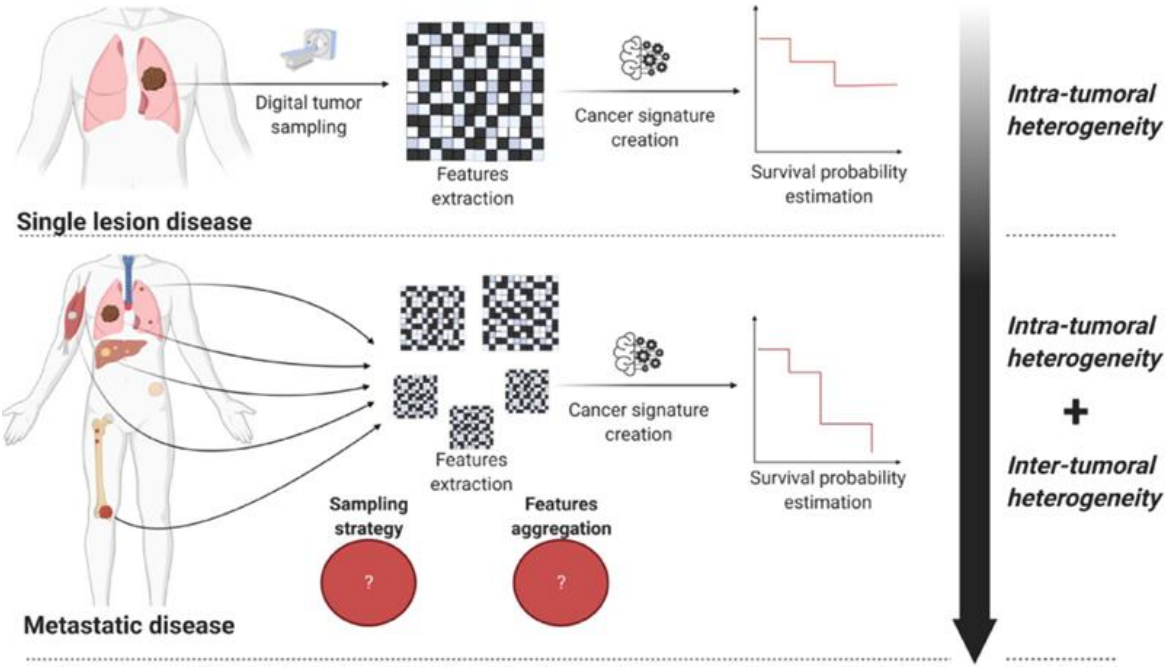

Figure 1: Comparison of radiomic analysis pipelines in single lesion and metastatic disease (reused from T. Henry et al, This is an open access article distributed under the terms of the Creative Commons CC BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited [47].

4.2 AI-Powered Protocol Design and Adaptive Learning Systems

The inflexibility of conventional clinical trial methods restricts their capacity to adjust to changing patient reactions and new biomarker discoveries. By facilitating the creation of real-time, adaptive trial methods, AI gets around this restriction. One example of this innovation is Trial Pathfinder, which simulated and optimized trial eligibility criteria using real-world data from more than 61,000 cancer patients. This system demonstrated how AI-designed procedures enhanced trial effectiveness and survival rates [52]. Bayesian adaptive designs are also supported by AI, which modifies dose levels and patient randomization in response to interim results. For example, AI was used to dynamically allocate patients with breast cancer to neoadjuvant therapy in the I-SPY2 trial, which greatly increased trial response [53]. Dose escalation, cohort selection, and adaptive halting are just a few of the trial flow optimization techniques being used by reinforcement learning, a subfield of machine learning in which computers learn by trial and error [54]. The following examples demonstrate the use of adaptive protocols in precision oncology, and they are included in Table 1.

Table 1: Innovative Trial Designs in Precision Oncology

Trial Type / Study	**Key Features**	**Benefits**	**Limitations**
N-of-1 Trials	Personalized treatment for each patient based on molecular profile. Comparisons made to historical/real-world data.	Tailored therapy for complex, heterogeneous tumors.	No standard comparator; complex data analysis; treatment variability between patients.
I-PREDICT	Multidisciplinary trial using tumor profiling, ctDNA, PD-L1, TMB, MSI to create a matching score guiding combo therapy.	Higher matching score : better disease control, PFS, OS.	Complex logistics; requires multidisciplinary coordination and deep molecular insights.
WINTHER Trial	Patients matched to therapy via genomics (Arm A) or transcriptomics (Arm B). PFS2	RNA and DNA profiling help improve treatment matching.	Did not meet primary endpoint; requires large-

	compared to PFS1 (Von Hoff model).		scale profiling infrastructure.
Home-Based Trials	Patients receive treatment and monitoring at home. Uses digital health tools and mobile nurses.	Increases access and recruitment, esp. in remote areas; reduces burden on infrastructure.	Challenges in monitoring adverse events and treatment response in real time.
Alpha-T Trial (Home-Based)	Evaluates alectinib in rare ALK+ solid tumors via a phase II, tissue-agnostic, single-arm home-based design.	Reaches ultra-rare cancer populations; improves trial inclusivity.	Still in progress; results pending; logistical coordination with mobile care needed.
Just-in-Time Activation	Sites activated rapidly once a matching patient is found. Useful for rare genotypes.	Speeds up trial access for rare cases.	Site setup delays still possible; early patient identification not always feasible.

4.3 Integration of Multimodal Data: From Genomics to Real-World Evidence

The current era of precision oncology also calls for the convergence of multi-modal data, including genomics, transcriptomics, proteomics, imaging, digital health data, and patient-reported outcomes, to personalize clinical trial design and treatment approaches. Figure 2 represents the paradigm of personalized medicine, in which treatment approaches are informed by genetic, environment, and behavioral factors [55]. The convergence of multiple layers of information is paving the way for the shift from disease-centric to mutation-centric clinical trial design. AI-driven solutions such as MOGONET and DeepOmix demonstrate the utility of the convergence of multi-omic information with the application of graph neuronal networks and deep learning approaches to classify subtypes of diseases, predict response, and stratify patients [56]. Basket, umbrella, and platform trials (Table 2) in clinical trials intend to demonstrate the application of convergence in practice, in which AI promotes real-time patient matching and adaptive arm switching. Moreover, real-world data sources such as electronic health record platforms, wearables, and health registries (Table 3) also provide insights with higher accuracy compared to current clinical trial cohorts. AI-driven multimodal approaches also facilitate the application of real-world data in regulatory settings (RWE) to support new approved drugs and post-market surveillance [57]. However, there remain certain limitations, including the need to achieve data standardization, AI model interpretability, preservation, and data exchange in

multiple institutions, which need to be overcome in achieving the application of multimodal AI in clinical trials.

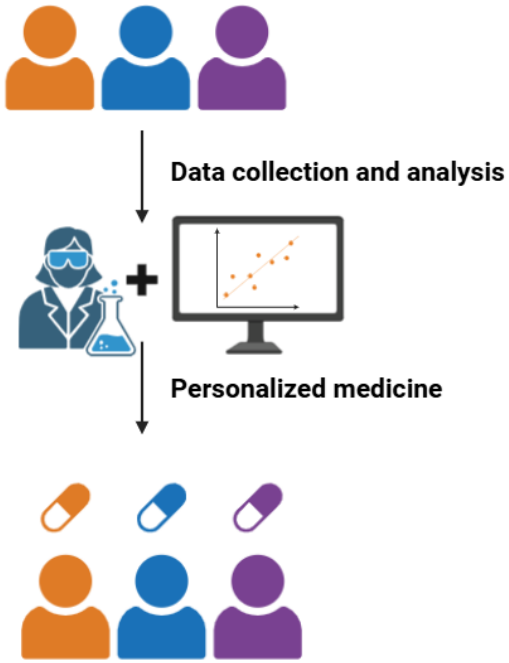

Figure 2: Personalized medicine framework showing the integration of genetic, environmental, and lifestyle variables (designed for this review).

Table 2: Data-Driven Trial Designs in Precision Oncology.

Trial Design	**Definition**	**Representative Trials**	**Key Features**	**Challenges**
Basket Trials	Test a targeted therapy for a specific mutation across different tumor types	-KEYNOTE (Pembrolizumab) -LOXO-TRK, NAVIGATE (Larotrectinib, Entrectinib)	-Tumor-agnostic -Gene-specific -FDA approvals for MSI-H, TMB-H, NTRK fusions	-Tumor heterogeneity. -Resistance mechanisms -Rare mutations

Umbrella Trials	Test multiple therapies in one tumor type based on different biomarkers	- Lung-MAP (NSCLC) - ALCHEMIST -I-SPY2 (breast) -plasmaMATCH	-Single histology - Multi-arm, biomarker-driven - Molecular stratification	-Biomarker assay complexity -Low efficacy for some matches -Rare subgroups
Platform Trials	Evaluate multiple hypotheses/therapies under one protocol with adaptive design	-IMPACT1 & 2 -TAPUR -NCI-MATCH -STAMPEDE -DART	-Adaptive arms (add/drop based on results) -Across multiple tumor types or one type -Real-world integration	-Complex logistics/statistics - Long follow-up -Cost and heterogeneity management

Table 3: Novel Mechanisms of Data Collection

Mechanism	**Description**	**Benefits**	**Challenges/ Limitations**
Exceptional Responders	Analyze rare patients with unusually strong responses to treatment using comprehensive tumor sequencing to identify predictive mutations.	Identify strong predictive biomarkers, understand drug mechanisms, reduce trial size and cost.	Small sample size, limited data harmonization, difficulty linking clinical features to outcomes.
Registry Protocols	Use structured clinical registries with demographic, treatment, and biologic data across large populations.	Provide real-world insights, reduce trial cost, allow broad evaluation of drug effectiveness.	Data may lack clinical precision, hard to collect and analyze timely data, less controlled than RCTs.

Real-World Data (RWD)	Data from EHRs, digital apps, insurance claims, or observational databases used to assess drug safety and effectiveness outside clinical trials.	Include underrepresented populations, accelerate approvals, broader safety/efficacy assessment.	Documentation errors, data heterogeneity, difficult standardization and interpretation.
Patient-Reported Outcome Measures (PROMs)	Data directly from patients via platforms/apps about symptoms, side effects, and quality of life during trials.	Enhance symptom control, improve survival, reduce ER visits, increase quality of care.	Expensive to implement, digital literacy issues, less precise self-reporting, rarely accepted by regulators.

4.4 Ethical, Regulatory, and Operational Challenges

Although AI presents immense transformative potential in the design of precision oncology trials, the successful implementation of AI is dependent on overcoming certain ethical, regulatory, and operational obstacles. These issues arise due to the complexities in the application of AI in the health care systems, along with the nature of the data.

A. Ethical Challenges: Bias, Explainability, and Informed Consent

Algorithmic bias ranks among the most pressing issues in relation to ethics, inasmuch as AI models rely on homogenous and/or incomplete data, causing inequitable outcomes. For example, ethnic minorities' data incompleteness might result in suboptimal recommendations or ineligibility for trials [58]. Moreover, “most successful AI models rely on ‘black box’ methodologies, which makes tracking predictions even more complex for clinicians and patients” [59]. This matters in terms of understanding clinical perceptions and informed consent in cases when AI models partake in trials and treatment. Emerging alternatives to AI models, termed Explainable AI (XAI) and continuous consent models with digital platforms, have been proposed with the goal of producing interpretable results while sustaining performance capacity, though their application remains to be seen [59]. AI models in clinical trials require transparency. On the one hand, feature weights in embedded models might be interpreted. However, complex models such as deep learning require post-hoc solutions such

as SHAP, LIME, and permutation importances. On the other, actions taken by AI models might be clarified in terms of what change in input might cause an altered outcome in classification. Explanations must then be tracked in the clinical environment to ascertain they meet trial thresholds in terms of triage, fairness, and drift [60], [61].

B. Regulatory Challenges: Lack of Standard Frameworks

There exists an evolving landscape related to the application of AI in designing clinical trials and developing medical devices. Although the FDA has provided a proposed regulatory approach to AI/ML in Software as a Medical Device (SaMD) guidance, it is being further shaped [62]. The European Medicines Agency (EMA) released its first draft of the guideline related to the usage of AI in the life cycle of medicinal products in 2023. The need for human involvement, model validation, and transparency was reinforced in the guideline [63]. Until globally harmonized rules and standards, the responsibility is upon the sponsor & investigators to comply with AI in clinical trials.

C. Operational Challenges: Data Sharing, Interoperability, and Deployment

AI systems require access to high-quality, multimodal datasets, yet data is frequently siloed across institutions due to legal, technical, or competitive barriers. Initiatives such as Swarm Learning and federated learning offer privacy-preserving solutions by enabling model training without centralized data sharing [64]. In addition, the lack of interoperability between hospital systems and trial platforms hinders smooth integration of AI tools. Standardizing data formats (HL7 FHIR), developing unified ontologies, and employing middleware APIs are necessary steps. Finally, AI tools require ongoing calibration and validation to ensure consistent performance across populations and over time-a process still absent in most trial infrastructure today. Addressing these ethical, regulatory, and operational challenges is vital to safely, fairly, and efficiently deploy AI technologies in oncology trials.

4.5 Toward a New Paradigm: Smart Trials for a Complex Era

The intersection of artificial intelligence, multi-omics data, digital health platforms, and decentralized clinical trial infrastructure is revolutionizing the future of clinical trials. This represents the dawn of a new age-SMART Trials, which stand for Smart Trials: Digital Trials, Adaptive, Intelligent, responsive to patient sub-populations, response variables, biomarkers, and even environmental/behavioral factors. While randomized controlled trials (RCTs) remain the hallmark of clinical trials, SMART Trials incorporate real-time analytics, federated data

infrastructures, and wearables to enhance flexibility, diversity, and efficiency. These trials adapt not only to patient sub-populations, but also to evolving clinical data, emerging biomarkers, and even environmental and behavioral factors. One fundamental innovative concept driving the change in clinical trials is the concept of DIGITAL TWINS. Digital twins refer to computational models of patients/cohorts, which forecast disease evolution, treatment response, and treatment risk, with data input parameters specific to the patient/cohorts. Digital twins facilitate the optimization of protocol limbs, pre-trial forecasts, and prospective analysis of treatments in clinical trials, entirely without any risk to patient safety [65]. Another innovative paradigm emerging in the clinical trials environment is FEDERATED LEARNING. This refers to machine learning in decentralized infrastructures, without the movement and storage of patient data, accumulated in central databases. Thus, continuous learning with collaborations between multiple data points in trials situated in territories globally, with preservation of patient privacy, becomes possible [64]. Moreover, mobile health solutions, such as sensors, eConsent solutions, and AI-enabled symptom reporting apps, reduce the burden on patients, including in resource-impoverished settings.

5. Institutional Case Studies in Translational Precision Oncology

Even if precision oncology were a global affair, the history of precision oncology has been shaped in ways that have been disproportionately led by certain institutions, such as Harvard University, the Massachusetts Institute of Technology (MIT), Oxford University, and Cambridge University. These play the role of hubs of innovation, not just in terms of advancing our understanding of the genome, but also in terms of developing clinical, computational, and ethical frameworks with which we redefine the practice of cancer science.

5.1 Harvard University and Dana-Farber Cancer Institute

Harvard's matrix of exemplary hospitals and research institutions, including Dana-Farber Cancer Institute (DFCI), Brigham and Women's Hospital, and Massachusetts General Hospital, is interwoven with translational oncology. These institutions are the co-leaders of the Profile Project, one of the world's largest institutional clinical sequencing initiatives. The project has generated data on more than 35,000 cancer patients and established mutation prevalence benchmarks across various tumor types [66]. Harvard also leads NCI-MATCH, which is a

tumor-agnostic trial that matches patients to targeted therapies based on next-generation sequencing (NGS)-based molecular alterations and promising therapy in tumor agnostic indications. This set the stage to create the biomarker-first stratification frameworks that are now used worldwide. Further, the creation of MatchMiner, an open-source artificial intelligence-enabled clinical trial matching platform, is driving trial enrolment and increasing equity of access [8]. In addition, Harvard works with Flatiron Health, Tempus, and Foundation Medicine to integrate real world evidence (RWE) into prospective trial planning. This partnership produces regulatory-grade observational analyses, which are now accepted by the FDA as acceptable supportive evidence for label expansions and drug repurposing decisions [67]. Moreover, tumor boards that are affiliated with Harvard increasingly rely upon multi-omic dashboards that are enabled by explainable AI. This allows oncologists to use transcriptomics, radiomics, and proteomics in additive ways to inform real-time treatment decisions.

5.2 Massachusetts Institute of Technology (MIT)

MIT's unique potential resides in their integrated take on engineering, AI, and biomedical science. For example, the Koch Institute was responsible for a number of advancements including tumour-targeted nanoparticles, programmable drug delivery, and synthetic biology diagnostics; They are even in the preclinical stage for potential early diagnosis and stratification of trials technologies, such as CRISPR-Cas sensors for detecting circulating tumour DNA [64].

The Jameel Clinic (referred to as J-Clinic) collaborates with graph neural networks (GNNs) and deep reinforcement learning to provide models of virtual patients, design adaptive trials, and predict adverse event risks before the patient receives treatment [12], [23]. These technologies also allow optimising the dose schedule, mitigate the risk of dropping out of a study after randomisation and predict drugs's potentially synergistic interactions. With support from the FDA Oncology Center of Excellence, MIT's digital twin program gives researchers the ability to conduct in silico clinical trials. This approach decreases the likelihood of amending research protocols and reduces the time spent in regulatory review. Their recent work using digital pathology, patient-reported outcomes, as well as, calibrating twins with time-series imaging, such as [65].

5.3 University of Oxford

Oxford's leadership in genomic infrastructure, data governance, and machine learning is encompassed in a portfolio of institutes e.g. Big Data Institute, Wellcome Centre for Human Genetics and Department of Oncology. Oxford is co-leading the 100,000 Genomes Project, where it has innovated somatic variant detection, structural variation and clinical grade reporting pipelines [68], [22]. In addition, its contribution to the GenOMICC study has allowed precision genomics to be applied to cases of rare, aggressive and treatment resistant cancers, especially ones that affect minority ethnic populations [22]. Its involvement in FOCUS4 - a genomically stratified adaptive trial in colorectal cancer highlights Oxford's capacity to explore and bring to the regulatory grade translational research. Further, Oxford's work in PAN-COVID has provided definitions of care models for cancer patients during pandemic scenarios and impacted practice and policy at all levels of the NHS. Ethics and regulation is central to Oxford's purpose. Oxford has collaborated with MHRA, EMA and GA4GH to shape draft guidance for AI explainability, data portability and model monitoring when genomics is applied in clinical care.

5.4 University of Cambridge

The Cambridge site combines discovery-level biology with implementation-level clinical translation, and has strong backing from the CRUK Cambridge Institute, CRUK RadNet, and Cambridge University Hospitals NHS Foundation Trust. The site is a founding site of the TRACERx study, which performs longitudinal ctDNA profiling, single-cell sequencing, and multiregional biopsies to visualize clonal dynamics in individuals with early-stage NSCLC Cambridge has provided data using liquid biopsy endpoints to create adaptive cohorts in the neoadjuvant and post-operative context [69]. The RadNet programme develops and tests innovations in radiogenomics, spatial transcriptomics, and immune landscape characterizations to understand the predictors of radioresistance. Cambridge is also validating the use of exosomal RNA and fragmentomics as a surrogate trial endpoint through the CAPTURE and SIGNATURE studies, technologies that are being adopted as part of pan-European biomarker consortia [70]. Cambridge is also at the forefront of AI-powered clinical platforms. Its ongoing work on automated tumor boards, which merges imaging, histopathology, and genomics via multimodal neural networks is undergoing evaluation by ESMO and the UK NHS Cancer Alliances for potential deployment nationally [30]. Beyond the academic contributions, the

active partnership with AstraZeneca, Illumina, and GRAIL on the Cambridge Biomedical Campus is contributing towards breaking down the wall between discovery and implementation science, and enabling only-then-phased first-in-human trials.

5.5 Synthesis

Every institution offers something translatable to our shared mission of scalable, equitable, and ethically responsible translational oncology in unique ways - Harvard with its AI-powered clinical systems, MIT has programmable AI and digital twins, Oxford has federated data governance, and Cambridge has multi-modal discovery and implementation of biomarkers. Together these institutions demonstrate a 'living' blueprint of how academic, clinical, and regulatory institutions can co-evolve together to provide radically personalised cancer care at unprecedented scale.

6. Future Perspectives

The evolution of precision oncology is no longer aspirational but now operational. As shown through institutional case studies and the technological pathways we have mapped, genomics, AI, and the innovation of trials are radically changing how we understand and treat cancer. However, for the change to be equitable, scalable, and sustainable, we must achieve advancement synchronously across three converging dimensions: the scientific infrastructure, regulatory capacity, and political alignment.

6.1 Technological and Scientific Horizons

Future trials will establish liquid biopsies in real time, utilize digital twins, and, with the growing power of multi-omics approaches and AI-based protocol updating, will utilize synthetic control arms instead of traditional comparators which offer both increased efficiency as well as greater ethical concordance [71]. With the ever-increasing portrayal of spatial omics, quantum inspired algorithms and long-read sequencing will offer greater dimension into tumour microenvironments, immune niches and rare subclonal populations [72]. Interoperability will depend on platforms in the cloud, and the utilization of federated learning which as I discussed above will require compliance with global data standards (e.g. HL7 FHIR, GA4GH, ISO/IEC 27001). Recent progress in medical AI has emphasized parameter-efficient learning and vision–language integration. Qin et al. demonstrated that “frozen-backbone” adapters can preserve prior medical knowledge while cutting trainable parameters by more than

90%, enabling robust transfer across domains [73]. Liu et al. developed a Global-to-Dense (G2D) radiography pre-training framework, combining global context with dense feature prediction to improve fine-grained clinical interpretation [74]. These medical AI advances exemplify scalable strategies that can be adapted to genomics-driven oncology, where efficient models and multimodal fusion are critical for real-world translation.

6.2 Regulatory Convergence and Global Health Equity

Regulatory bodies will have to shift toward harmonised frameworks for AI diagnostics and adaptive paths. Initiatives such as the FDA's SaMD Action Plan, EMA's AI Guidance, and the IMDRF AI Working Group should synchronise to mitigate fragmentation, and facilitate approvals [75]. Equity means inclusion - inclusion of diverse genomic ancestries and also inclusion of low- and middle-income countries (LMICs) in trial design and genomic infrastructure. Programs such as All of Us, GenOMICC, and the African Genome Project offer ethical paths forward [76].

6.2.1 Equity and dataset representativeness

Under-representation of minority and low-resource populations risks biased predictions and poor calibration. Remedies include governance mandates for subgroup reporting, privacy-preserving federated or swarm learning to expand datasets, domain adaptation to correct imbalance, and community genomics programs that reinvest locally. Journals and regulators should require subgroup AUC and calibration plots to ensure equitable generalizability [77].

6.3 Political Science, Policy, and the Governance of Innovation

Genomic medicine is, at least in equal parts, a political project as much as it is a scientific project [78]. Policy frameworks will need to consider [79]:

- data ownership and nationalism
- AI validation across jurisdictions
- collaboration of public-private entities
- confidence in regulatory bodies

The COVID-19 pandemic brought to light issues in global governance that are tenuous at best, and the urgency applied to new cancer policy should be prioritized similarly given new pandemic patterns, the vaccine-cancer immunotherapy nexus, and antimicrobial resistance [80]. The two universities Oxford and Cambridge’s involvement in health diplomacy through

WHO and UNESCO, and MIT's participation with DARPA in developing translational pathways, reiterate the new geopolitical role research institutions will have, in the governance of cancers globally.

Offering "Genomics-Driven Precision Oncology as a Service" creates confidentiality challenges. Privacy-by-design approaches-federated or swarm learning, secure aggregation, and standardized data-use agreements-can enable scalability without raw data transfer. Alignment with the EU AI Act and FDA SaMD guidelines is essential. Deployments should prioritize sovereign-cloud or on-premises solutions, supported by Data Protection Impact Assessments (DPIAs) that summarize residual risks and safeguards [81], [82].

6.4 **Final Reflections**

We are at a critical juncture. Precision oncology is no longer restricted to elite cancer institutions; it is increasingly accessible at community hospitals, developing countries, and via virtual platforms. The determination of whether precision oncology is a universal right or only a distinct privilege will hinge on the integration of science, data, and ethics, bolstered by inclusive institutions, policy foresight, and civic trust.

7. **Discussion and Conclusion**

In conclusion, this paper has presented the revolutionary change brought by the junction of genetic innovation, adaptive trials, and institutional developments in the realm of precision oncology. At the core, this revolution represents the application of AI, real-world evidence, and HTS to transform biomarker-based approaches from rigid procedural routines in the realm of precision oncology. Case studies involving Harvard, MIT, Oxford, and Cambridge universities represent the science acuity and translational foresight in integrating 'omics' data in clinical routines. However, there persist important hindrances in the path. Rather, the current models in precision oncology in terms of generalizability persist in facing inequities in patient enrollment, variant misclassification, and regulatory discordance. These points, specifically, serve as current limitations. Thus, the need arises to incorporate flexibility in future structuring in strongly important ethical, regulatory, and policy considerations. There exists possible benefit in global regulatory structures with inclusive, transparent, and accountable data infrastructures. Rather, precision oncology, with initiation in theory in select premier institutions, represents the current practical application in leader institutions in the world, although one surmises personal conviction in assuming major global promise in resolution.

This paper presents the innervating role played in inclusive design, ethical, and infrastructural necessity in concert with scientific advancement. Rather, with proper usage and global availability, the major promise of emerging innovative frontiers such as AI, liquid biopsy, and single-cell genomics will be realized. Broadly, the integrity to meet global regulatory advancement in terms of need, public belief, and productive structuring will serve in major shaping in the treatment options in cancers in the future. Rather, with convergence of science in treatment, precision-to-global unanimity represents the major promise in attaining global recognition in mainstream treatment options in cancers.

List of Abbreviations:

- NCI-MATCH: National Cancer Institute – Molecular Analysis for Therapy Choice trial
- ASCO TAPUR: American Society of Clinical Oncology – Targeted Agent and Profiling Utilization Registry
- TCGA The Cancer Genome Atlas
- CRISPR Clustered Regularly Interspaced Short Palindromic Repeats
- NGS Next-generation Sequencing
- ICGC International Cancer Genome Consortium
- MSK-IMPACT Memorial Sloan Kettering–Integrated Mutation Profiling of Actionable Cancer Targets

Author Contributions

Conceptualization: R.M.A, Original draft preparation: A.S, Writing-review and editing: R.M.A and A.S, Figure design, and visualization: R.M.A, Supervision: R.M.A, Validation, and critical revision: R.M.A and A.S. All authors have read and approved the final version of the manuscript.

Funding

No external funding was received for this study.

Conflicts of Interest

The authors declare no conflicts of interest regarding this manuscript.

Acknowledgement

The authors would like to acknowledge that OpenAI's ChatGPT was used for grammar check and DALL·E model was used through ChatGPT to generate the graphical abstract. The authors made sure all information are accurate.

References

[1] J. Zhang *et al.*, "International cancer genome consortium data portal-a one-stop shop for cancer genomics data," *Database*, vol. 2011, 2011, doi: 10.1093/DATABASE/BAR026,.

[2] The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 490, 61–70 (2012). https://doi.org/10.1038/nature11412

[3] G. Siravegna, S. Marsoni, S. Siena, and A. Bardelli, "Integrating liquid biopsies into the management of cancer," *Nat Rev Clin Oncol*, vol. 14, no. 9, pp. 531–548, Sep. 2017, doi: 10.1038/NRCLINONC.2017.14,.

[4] M. F. Mosele *et al.*, "Recommendations for the use of next-generation sequencing (NGS) for patients with advanced cancer in 2024: a report from the ESMO Precision Medicine Working Group," *Annals of Oncology*, vol. 35, no. 7, pp. 588–606, Jul. 2024, doi: 10.1016/J.ANNONC.2024.04.005.

[5] D. E. Spratt *et al.*, "Racial/Ethnic Disparities in Genomic Sequencing," *JAMA Oncol*, vol. 2, no. 8, pp. 1070–1074, Aug. 2016, doi: 10.1001/JAMAONCOL.2016.1854.

[6] A. Sveen, S. Kopetz, and R. A. Lothe, "Biomarker-guided therapy for colorectal cancer: strength in complexity," *Nat Rev Clin Oncol*, vol. 17, no. 1, p. 11, Jan. 2019, doi: 10.1038/S41571-019-0241-1.

[7] M. Gerlinger *et al.*, "Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing," *New England Journal of Medicine*, vol. 366, no. 10, pp. 883–892, Mar. 2012, doi: 10.1056/NEJMOA1113205,.

[8] K. T. Flaherty *et al.*, "The molecular analysis for therapy choice (NCI-MATCH) trial: Lessons for genomic trial design," *J Natl Cancer Inst*, vol. 112, no. 10, pp. 1021–1029, Oct. 2020, doi: 10.1093/JNCI/DJZ245,.

[9] A. Zehir *et al.*, "Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients," *Nat Med*, vol. 23, no. 6, pp. 703–713, Jun. 2017, doi: 10.1038/NM.4333,.

[10] M. W. Tibbitt, J. E. Dahlman, and R. Langer, "Emerging Frontiers in Drug Delivery," *J Am Chem Soc*, vol. 138, no. 3, pp. 704–717, Jan. 2016, doi: 10.1021/JACS.5B09974,.

[11] Fda, Cder, Bent, and Robyn, "Adaptive Designs for Clinical Trials of Drugs and Biologics Guidance for Industry," 2019, Accessed: Jul. 28, 2025. [Online]. Available:

[12] M. Jamal-Hanjani *et al.*, "Tracking the Evolution of Non–Small-Cell Lung Cancer," *New England Journal of Medicine*, vol. 376, no. 22, pp. 2109–2121, Jun. 2017, doi: 10.1056/NEJMOA1616288,.

[13] Z. Zhang *et al.*, "Genetic analyses support the contribution of mRNA N 6-methyladenosine (m6A) modification to human disease heritability," *Nat Genet*, vol. 52, no. 9, pp. 939–949, Sep. 2020, doi: 10.1038/S41588-020-0644-Z,.

[14] F. Iorio *et al.*, "A Landscape of Pharmacogenomic Interactions in Cancer," *Cell*, vol. 166, no. 3, pp. 740–754, Jul. 2016, doi: 10.1016/J.CELL.2016.06.017,.

[15] J. Barretina *et al.*, "The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity," *Nature 2012 483:7391*, vol. 483, no. 7391, pp. 603–607, Mar. 2012, doi: 10.1038/nature11003.

[16] L. B. Alexandrov *et al.*, "Signatures of mutational processes in human cancer," *Nature*, vol. 500, no. 7463, pp. 415–421, Aug. 2013, doi: 10.1038/nature12477

[17] A. Roth *et al.*, "PyClone: Statistical inference of clonal population structure in cancer," *Nat Methods*, vol. 11, no. 4, pp. 396–398, Mar. 2014, doi: https://doi.org/10.1038/nmeth.2883

[18] O. Shalem *et al.*, "Genome-scale CRISPR-Cas9 knockout screening in human cells," *Science (1979)*, vol. 343, no. 6166, pp. 84–87, 2014, doi: https://doi.org/10.1126/science.1247005

[19] J. M. Dempster *et al.*, "Agreement between two large pan-cancer CRISPR-Cas9 gene dependency data sets," *Nature Communications 2019 10:1*, vol. 10, no. 1, pp. 1–14, Dec. 2019, doi: 10.1038/s41467-019-13805-y.

[20] B. Adamson *et al.*, "A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response," *Cell*, vol. 167, no. 7, pp. 1867-1882.e21, Dec. 2016, doi: 10.1016/j.cell.2016.11.048.

[21] N. Sachs *et al.*, "A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity," *Cell*, vol. 172, no. 1–2, pp. 373-386.e10, Jan. 2018, doi: 10.1016/j.cell.2017.11.010.

[22] C. Turnbull *et al.*, "The 100 000 Genomes Project: bringing whole genome sequencing to the NHS," *BMJ*, vol. 361, Apr. 2018, doi: 10.1136/BMJ.K1687.

[23] R. Guo, J. Luo, J. Chang, N. Rekhtman, M. Arcila, and A. Drilon, "MET-dependent solid tumours — molecular diagnosis and targeted therapy," *Nat Rev Clin Oncol*, vol. 17, no. 9, pp. 569–587, Sep. 2020, doi: 10.1038/S41571-020-0377-Z,.

[24] H. Gao *et al.*, "The landscape of tolerated genetic variation in humans and primates," *Science*, vol. 380, no. 6648, p. eabn8153, Jun. 2023, doi: 10.1126/science.abn8197.

[25] J. Cheng *et al.*, "Accurate proteome-wide missense variant effect prediction with AlphaMissense," *Science*, vol. 381, no. 6664, Sep. 2023, doi: https://doi.org/10.1126/science.adg7492

[26] G. M. Findlay *et al.*, "Accurate classification of BRCA1 variants with saturation genome editing," *Nature*, vol. 562, no. 7726, pp. 217–222, Oct. 2018, doi: https://doi.org/10.1038/s41586-018-0461-z

[27] T. M. Gierahn *et al.*, "Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput," *Nature Methods 2017 14:4*, vol. 14, no. 4, pp. 395–398, Feb. 2017, doi: 10.1038/nmeth.4179.

[28] H. O. Lee *et al.*, "Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer.," *Nat Genet*, vol. 52, no. 6, pp. 594–603, May 2020, doi: 10.1038/S41588-020-0636-Z.

[29] X. Wu, F. Mu, Y. Wang, and H. Zhao, "Application of atomic simulation methods on the study of graphene nanostructure fabrication by particle beam irradiation: A review," *Comput Mater Sci*, vol. 149, pp. 98–106, Jun. 2018, doi: 10.1016/j.commatsci.2018.03.022.

[30] F. Mosele *et al.*, "Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: a report from the ESMO Precision Medicine Working Group," *Annals of Oncology*, vol. 31, no. 11, pp. 1491–1505, Nov. 2020, doi: 10.1016/j.annonc.2020.07.014.

[31] Y. Yang *et al.*, "Efficacy and Safety of Sintilimab Plus Pemetrexed and Platinum as First-Line Treatment for Locally Advanced or Metastatic Nonsquamous NSCLC: a Randomized, Double-Blind, Phase 3 Study (Oncology pRogram by InnovENT anti-PD-1-11)," *Journal of Thoracic Oncology*, vol. 15, no. 10, pp. 1636–1646, Oct. 2020, doi: 10.1016/j.jtho.2020.07.014.

[32] S. C. Ng *et al.*, "Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies," *The Lancet*, vol. 390, no. 10114, pp. 2769–2778, Dec. 2017, doi: 10.1016/S0140-6736(17)32448-0.

[33] M. Ligero, O. S. M. El Nahhas, M. Aldea, and J. N. Kather, "Artificial intelligence-based biomarkers for treatment decisions in oncology," *Trends Cancer*, vol. 11, no. 3, pp. 232–244, Mar. 2025, doi: 10.1016/j.trecan.2024.12.001.

[34] R. Sisk *et al.*, "Informative presence and observation in routine health data: A review of methodology for clinical risk prediction," *J Am Med Inform Assoc*, vol. 28, no. 1, p. 155, Jan. 2021, doi: 10.1093/JAMIA/OCAA242.

[35] K. S. Thress *et al.*, "Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M," *Nat Med*, vol. 21, no. 6, pp. 560–562, Jun. 2015, doi: 10.1038/NM.3854,.

[36] L. Horn *et al.*, "First-Line Atezolizumab plus Chemotherapy in Extensive-Stage Small-Cell Lung Cancer," *New England Journal of Medicine*, vol. 379, no. 23, pp. 2220–2229, Dec. 2018, doi: 10.1056/NEJMOA1809064/SUPPL_FILE/NEJMOA1809064_DATA-SHARING.PDF.

[37] J. Tie *et al.*, "Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer," *Sci Transl Med*, vol. 8, no. 346, Jul. 2016, doi: 10.1126/scitranslmed.aaf6219,.

[38] Y. Zhao, M. R. Kosorok, and D. Zeng, "Reinforcement learning design for cancer clinical trials," *Stat Med*, vol. 28, no. 26, pp. 3294–3315, Nov. 2009, doi: 10.1002/sim.3720,.

[39] K. B. Johnson *et al.*, "Precision Medicine, AI, and the Future of Personalized Health Care," *Clin Transl Sci*, vol. 14, no. 1, pp. 86–93, Jan. 2021, doi: 10.1111/cts.12884,.

[40] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, Mar. 2011, doi: 10.1016/j.cell.2011.02.013.

[41] E. Farina, J. J. Nabhen, M. I. Dacoregio, F. Batalini, and F. Y. Moraes, "An Overview of Artificial Intelligence in Oncology," *Future Sci OA*, vol. 8, no. 4, p. 787, Apr. 2022, doi: 10.2144/fsoa-2021-0074.

[42] C. Luchini, A. Pea, and A. Scarpa, "Artificial intelligence in oncology: current applications and future perspectives," *Br J Cancer*, vol. 126, no. 1, pp. 4–9, Jan. 2022, doi: 10.1038/S41416-021-01633-1

[43] R. Hamamoto *et al.*, "Application of artificial intelligence technology in oncology: Towards the establishment of precision medicine," *Cancers (Basel)*, vol. 12, no. 12, pp. 1–32, Dec. 2020, doi: 10.3390/cancers12123532,.

[44] L. Kolla and R. B. Parikh, "Uses and limitations of artificial intelligence for oncology," *Cancer*, vol. 130, no. 12, pp. 2101–2107, Jun. 2024, doi: 10.1002/CNCR.35307,.

[45] S. Ramón y Cajal *et al.*, "Clinical implications of intratumor heterogeneity: challenges and opportunities," *J Mol Med*, vol. 98, no. 2, pp. 161–177, Feb. 2020, doi: 10.1007/S00109-020-01874-2.

[46] T. Miyauchi, T. Yaguchi, and Y. Kawakami, "Inter-patient and intra-tumor heterogeneity in the sensitivity to tumor-targeted immunity in colorectal cancer," *Japanese Journal of Clinical Immunology*, vol. 40, no. 1, pp. 54–59, 2017, doi: 10.2177/JSCI.40.54,.

[47] T. Henry *et al.*, "Investigation of radiomics based intra-patient inter-tumor heterogeneity and the impact of tumor subsampling strategies," *Sci Rep*, vol. 12, no. 1, pp. 1–11, Dec. 2022, doi: 10.1038/S41598-022-20931-Z

[48] R. Miotto, L. Li, B. A. Kidd, and J. T. Dudley, "Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records," *Sci Rep*, vol. 6, no. 1, pp. 1–10, May 2016, doi: 10.1038/srep26094

[49] J. K. Sicklick *et al.*, "Molecular profiling of cancer patients enables personalized combination therapy: the I-PREDICT study," *Nat Med*, vol. 25, no. 5, pp. 744–750, May 2019, doi: 10.1038/S41591-019-0407-5,.

[50] B. Koçak, E. Ş. Durmaz, E. Ateş, and Ö. Kılıçkesmez, "Radiomics with artificial intelligence: a practical guide for beginners," *Diagnostic and Interventional Radiology*, vol. 25, no. 6, p. 485, Nov. 2019, doi: 10.5152/dir.2019.19321.

[51] A. Zwanenburg *et al.*, "The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping," *Radiology*, vol. 295, no. 2, pp. 328–338, May 2020, doi: 10.1148/radiol.2020191145.

[52] R. Liu *et al.*, "Evaluating eligibility criteria of oncology trials using real-world data and AI," *Nature*, vol. 592, no. 7855, pp. 629–633, Apr. 2021, doi: 10.1038/S41586-021-03430-5

[53] H. Wang and D. Yee, "I-SPY 2: a Neoadjuvant Adaptive Clinical Trial Designed to Improve Outcomes in High-Risk Breast Cancer," *Curr Breast Cancer Rep*, vol. 11, no. 4, p. 303, Dec. 2019, doi: 10.1007/S12609-019-00334-2.

[54] B. T. Li *et al.*, "Reimagining patient-centric cancer clinical trials: a multi-stakeholder international coalition," *Nat Med*, vol. 28, no. 4, pp. 620–626, Apr. 2022, doi: 10.1038/S41591-022-01775-6.

[55] S. Parvizpour, H. Beyrampour-Basmenj, J. Razmara, F. Farhadi, and M. S. Shamsir, "Cancer treatment comes to age: from one-size-fits-all to next-generation sequencing (NGS) technologies," *Bioimpacts*, vol. 14, no. 4, p. 29957, Jul. 2023, doi: 10.34172/BI.2023.29957.

[56] T. Wang *et al.*, "MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification," *Nat Commun*, vol. 12, no. 1, pp. 1–13, Dec. 2021, doi: 10.1038/S41467-021-23774-W

[57] G. Alipour-Haris, X. Liu, V. Acha, A. G. Winterstein, and M. Burcu, "Real-world evidence to support regulatory submissions: A landscape review and assessment of use cases," *Clin Transl Sci*, vol. 17, no. 8, p. e13903, Aug. 2024, doi: 10.1111/CTS.13903.

[58] J. L. Cross, M. A. Choma, and J. A. Onofrey, "Bias in medical AI: Implications for clinical decision-making," *PLOS Digital Health*, vol. 3, no. 11, p. e0000651, Nov. 2024, doi: 10.1371/journal.pdig.0000651.

[59] Z. Sadeghi *et al.*, "A review of Explainable Artificial Intelligence in healthcare," *Computers and Electrical Engineering*, vol. 118, p. 109370, Aug. 2024, doi: 10.1016/j.compeleceng.2024.109370.

[60] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should i trust you?' Explaining the predictions of any classifier," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13-17-August-2016, pp. 1135–1144, Aug. 2016, doi: https://doi.org/10.1145/2939672.2939778

[61] S. M. Lundberg and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," *Adv Neural Inf Process Syst*, vol. 2017-December, pp. 4766–4775, May 2017, Accessed: Sep. 30, 2025. [Online]. Available: https://arxiv.org/pdf/1705.07874

[62] "Artificial Intelligence in Software as a Medical Device | FDA." Accessed: Jul. 28, 2025. [Online]. Available: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device

[63] "Committee for Medicinal Products for Human Use (CHMP) Committee for Medicinal Products for Veterinary Use (CVMP) Reflection paper on the use of Artificial Intelligence (AI) in the medicinal product lifecycle," 2024, Accessed: Jul. 28, 2025. [Online]. Available: www.ema.europa.eu/contact

[64] S. Warnat-Herresthal *et al.*, "Swarm Learning for decentralized and confidential clinical machine learning," *Nature 2021 594:7862*, vol. 594, no. 7862, pp. 265–270, May 2021, doi: 10.1038/s41586-021-03583-3.

[65] L. Mollica, C. Leli, F. Sottotetti, S. Quaglini, L. D. Locati, and S. Marceglia, "Digital twins: a new paradigm in oncology in the era of big data," *ESMO Real World Data*

and Digital Oncology, vol. 5, p. 100056, Sep. 2024, doi: 10.1016/J.ESMORW.2024.100056.

[66] E. M. Van Allen *et al.*, "Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine," *Nat Med*, vol. 20, no. 6, pp. 682–688, 2014, doi: 10.1038/nm.3559.

[67] R. E. Sherman *et al.*, "Real-World Evidence — What Is It and What Can It Tell Us?," *New England Journal of Medicine*, vol. 375, no. 23, pp. 2293–2297, Dec. 2016, doi: 10.1056/nejmsb1609216.

[68] E. Pairo-Castineira *et al.*, "Genetic mechanisms of critical illness in COVID-19," *Nature 2020 591:7848*, vol. 591, no. 7848, pp. 92–98, Dec. 2020, doi: 10.1038/s41586-020-03065-y

[69] "Tracking cancer evolution: the TRACERx Lung Study | Cancer Research UK." Accessed: Jul. 28, 2025. [Online]. Available: https://www.cancerresearchuk.org/funding-for-researchers/research-opportunities-in-harder-to-treat-cancers/tracking-cancer-evolution-the-tracerx-lung-study

[70] "RadNet – our radiation research network | Cancer Research UK." Accessed: Jul. 28, 2025. [Online]. Available: https://www.cancerresearchuk.org/funding-for-researchers/our-research-infrastructure/radnet-our-radiation-research-network

[71] Y. Jia *et al.*, "Transmembrane water-efflux rate measured by magnetic resonance imaging as a biomarker of the expression of aquaporin-4 in gliomas," *Nat Biomed Eng*, vol. 7, no. 3, pp. 236–252, Mar. 2023, doi: 10.1038/S41551-022-00960-9.

[72] V. Marx, "Method of the Year: spatially resolved transcriptomics," *Nat Methods*, vol. 18, no. 1, pp. 9–14, Jan. 2021, doi: 10.1038/S41592-020-01033-Y.

[73] J. Qin, C. Liu, S. Cheng, Y. Guo, and R. Arcucci, "FREEZE THE BACKBONES: A PARAMETER-EFFICIENT CONTRASTIVE APPROACH TO ROBUST MEDICAL VISION-LANGUAGE PRE-TRAINING," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 1686–1690, 2024, doi: 10.1109/ICASSP48485.2024.10447326.

[74] C. Liu, C. Ouyang, S. Cheng, A. Shah, W. Bai, and R. Arcucci, "G2D: From Global to Dense Radiography Representation Learning via Vision-Language Pre-training," *Adv Neural Inf Process Syst*, vol. 37, Dec. 2023, Accessed: Sep. 30, 2025. [Online]. Available: https://arxiv.org/pdf/2312.01522

[75] K. A. Khan and R. S. Kerbel, "Improving immunotherapy outcomes with anti-angiogenic treatments and vice versa," *Nat Rev Clin Oncol*, vol. 15, no. 5, pp. 310–324, May 2018, doi: 10.1038/nrclinonc.2018.9.

[76] P. Scalia, S. J. Williams, Y. Fujita-Yamaguchi, and A. Giordano, "Cell cycle control by the insulin-like growth factor signal: at the crossroad between cell growth and mitotic regulation," *Cell Cycle*, vol. 22, no. 1, p. 1, 2022, doi: https://doi.org/10.1080/15384101.2022.2108117

[77] J. Casaletto, A. Bernier, R. McDougall, and M. S. Cline, "Federated Analysis for Privacy-Preserving Data Sharing: A Technical and Legal Primer," *Annu Rev Genomics Hum Genet*, vol. 24, no. Volume 24, 2023, pp. 347–368, Aug. 2023, doi: 10.1146/annurev-genom-110122-084756/cite/refworks

[78] S. Beck, S. Jasanoff, A. Stirling, and C. Polzin, “The governance of sociotechnical transformations to sustainability,” *Curr Opin Environ Sustain*, vol. 49, pp. 143–152, Apr. 2021, doi: 10.1016/j.cosust.2021.04.010.

[79] C. Lyall and J. Tait, “Beyond the limits to governance: New rules of engagement for the tentative governance of the life sciences,” *Res Policy*, vol. 48, no. 5, pp. 1128–1137, Jun. 2019, doi: 10.1016/j.respol.2019.01.009.

[80] S. Taylor, “‘Global health’: Meaning what?,” *BMJ Glob Health*, vol. 3, no. 2, Mar. 2018, doi: 10.1136/bmjgh-2018-000843.

[81] “AI Act | Shaping Europe’s digital future.” Accessed: Sep. 30, 2025. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

[82] “Artificial Intelligence in Software as a Medical Device | FDA.” Accessed: Sep. 30, 2025. [Online]. Available: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device