Computing&AI Connect
ISSN: 3104-4719
2025, Article ID. 2025.0028, Cite as: https://doi.org/10.69709/CAIC.2025.144976

SCIFINITI PUBLISHING

Review Article

OPEN ACCESS

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

Arbiona Shukulli[✉,1] **Rahaf M. Ahmad**[✉,2,*]

[1] Preca College, Rruga Pandeli Vangjeli 5, Korçë 7001, Albania
[2] Department of Genetics and Genomics, College of Medicine and Health Sciences, United Arab Emirates University, Tawam Road, Al Maqam District, Al Ain, Abu Dhabi 15551, United Arab Emirates

Article History
Received: July 31, 2025 Accepted: December 15, 2025 Published: December 29, 2025

Abstract
Precision oncology has driven a paradigm shift in cancer care, from histology-guided regimens to genomics-informed, biomarker-driven therapies. This paper reviews major milestones in basic translational research in cancer genomics, advances in clinical trial design, and translational activities led by top Universities. This paper describes the major translational breakthroughs, such as TCGA, CRISPR-enabled functional genomics, and real-time trial-matching platforms (e.g., MatchMiner), and underscores the need for their collective advancement in stratifying patients and personalizing treatments. It also showcases how case studies from major institutions have aided the integration of multi-omics, adaptive clinical trials, and ethical AI approaches in research and clinical practice. Digital Twin models at 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, and AI-enriched clinical trials demonstrate an imminent future marked by learning health platforms. Despite major scientific advances, critical challenges remain in equity, data harmonization, and variant interpretation. Disparities in clinical trials and a lack of representation in diverse populations might accentuate global health inequities. Precision oncology brings together pioneering breakthroughs across technological, governance, political, and ethical domains. This paper also offers an inclusive blueprint for demarcating the limitations of Precision Oncology, intersecting with Genomic Understanding, innovative clinical trials, and Socio-politics, prying into pioneering 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

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 leverages genomic, transcriptomic, and epigenetic data to guide therapeutic decisions, the standards of care for many cancer types, including metastatic breast cancer, cholangiocarcinoma, neuroendocrine tumors and gliomas, have been redefined. Studies such as The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) have facilitated the transition from tumor classification based on histopathology to classification by 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, early diagnosis, and real-time monitoring by liquid biopsies [3,4]. Yet, although we are ushering in an era of precision oncology, the clinical translation of genomic knowledge into treatment remains chal-

* **Corresponding Author:**
Rahaf M. Ahmad, Department of Genetics and Genomics, College of Medicine and Health Sciences, United Arab Emirates University, Tawam Road, Al Maqam District, Al Ain, Abu Dhabi 15551, United Arab Emirates, 700039113@uaeu.ac.ae

lenging. Targeted agents have frequently failed in late-stage trials, and even successful therapies are often limited by acquired resistance. A persistent limitation is the underrepresentation of diverse and minority populations in genomic datasets and clinical trials [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 of the translational path of precision oncology, incorporating advances in genomics and the evolution of clinical trials.

Drawing on institutional insights from Harvard Medical School, the Broad Institute, MIT, the University of Oxford, and the University of Cambridge, this discussion defines how these hotbeds of excellence have driven breakthroughs in their respective genomics, while also establishing innovative approaches to clinical trial design. 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 to pioneering adaptive and biomarker-enhanced trial designs through programs such as the UK 100,000 Genomes Project, FOCUS4, and TRACERx [11,12]. It is argued that precision oncology should no longer be viewed as a linear process from mutation to drug. Rather, it should be considered a dynamic, iterative process that leverages real-time molecular information, AI-assisted methodologies, and patient-centered endpoints. The viability of systems under this framework depends not only on scientific advances but also on infrastructure, regulatory, and equity considerations. There is a need to overcome genomic data interpretation bottlenecks, harmonize international trial methods, and broaden access to molecular diagnostics, especially in low- and middle-income contexts. By synthesizing a broad range of data across multiple sub-disciplines and institutions, this review offers strategic and academic commentary on the historical implementation of genomics-led precision oncology through cancer clinical trials. Looking beyond current achievements, this study aims to identify and address factors that may impede the translation of genomic insights into impactful and measurable clinical outcomes, particularly in the context of pan-cancer, biomarker-driven, and AI-supported trials.

2. Genomic Foundations of Precision Oncology

Molecular characterization 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 disassembly of tumor biology with an unparalleled resolution, revealing layers of somatic mutations, copy number variation, structural rearrangements, chromatin organization, 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 characterization of a variety of tumor 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]. Together, these studies provided a springboard for pan-cancer analyses, enabling the identification of commonalities and differences among tumor types beyond their 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]. These datasets have helped not only to guide drug repurposing efforts, but also to develop and train machine learning models that predict therapeutic responses 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 tumors deficient in mismatch repair [16]. Cambridge has developed computational methods, including algorithms for clonal deconvolution (e.g., PyClone, SciClone), through the CRUK Cambridge Institute and the European Bioinformatics Institute (EBI). These algorithms have provided new insights con-

2025, Vol. 2, Article ID. 2025.0028
https://doi.org/10.69709/CAIC.2025.144976

cerning clonal architecture and evolutionary patterns as the patient is placed under treatment pressure [17]. Despite these advances, genomic datasets. While these assertions have improved generalizability across populations, limited availability remains a considerable barrier, including in TCGA, which is biased toward individuals of European ancestry. Even though these assertions have improved generalizability across populations, limited availability remains a substantial barrier to equitable clinical application. The underrepresentation of genomic analyses from diverse human populations is a significant issue for 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 enabled genome-wide studies of gene essentiality, identifying not only genetic dependencies but also synthetic lethal relationships (e.g., in BRCA-deficient cells perturbed with PARP inhibitors) and 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 characterize lineage-specific vulnerabilities or pan-cancer genetic dependencies. This map of cancer gene dependencies has already been used to design selective inhibitors targeting genes essential only in specific tumor types, such as WRN helicase in microsatellite instability-high (MSI-H) tumors and STAG2 in Ewing sarcoma [19]. At MIT, CRISPR screens have been adapted to integrate single-cell RNA sequencing readouts with pooled combinatorial perturbations to identify buffering networks and gain insights into gene–gene interactions, which are critically important for understanding 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 tumors in vitro [21]. Furthermore, AI-enabled models have recently begun combining 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 prioritization not only of 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 achieved 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 information, using platforms such as MatchMiner, which connect prospective patients with studies relevant to their circumstances [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 family risk assessment. The datasets also enable longitudinal EHRs, enabling correlative outcome analyses [22]. A major bottleneck is the high frequency of variants of uncertain significance (VUS), which complicates reporting and downstream clinical decision-making. 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, giving 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 tumor board scalability. The Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT and the Big Data Institute at Oxford are developing AI models to automate variant classification, prioritize targets, and predict response with multi-modal inputs. These products will be integrated into molecular tumor boards and the decision support interface.

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

Cancer exhibits pronounced spatial and temporal heterogeneity. Single cell RNA sequencing has enabled the identification of rare cellular subpopulations such as drug tolerant persisters and the reconstruction of lineage trajectories as cells transition through epithelial to mesenchymal transition states, including work conducted at the Broad Institute [27]. Spatial transcriptomics platforms like Slide-seq and 10× Genomics Visium are mapping the architectural relationships between cancer cells, stroma, and immune infiltrates. At Cambridge, spatial multi-omics analyses of colorectal and glioblastoma samples have identified immune exclusion zones and hypoxic niches that correlate with therapy 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) include scVI (single-cell variability inference) and DeepCell, which model integrated datasets comprising spatial transcriptomics, digital pathology, and single-cell epigenomics. These models can show how gene expression changes over time, how cells interact with one another, and how spatial architectures function. They represent a novel approach for stratifying clinical trials based on tumor topologies and microenvironmental characteristics.

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 have relied primarily on a histology-based stratification framework, which required tumors to be of a specific type and stage and did not account for the underlying molecular heterogeneity that drives 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 align with the molecular complexity of cancer, with trials quite broadly based on actionable mutations, without restrictions on tumor type (basket trials, such as NCI-MATCH and ASCO TAPUR). Harvard-affiliated centers such as Dana-Farber have contributed to basket-trial implementation by enrolling patients with different cancer types that share an actionable alteration [8,30]. Basket trials are designed to overcome the limitations of single-histology approaches, as oncogenic drivers can span multiple tissue origins, potentially enabling broader eligibility and faster patient accrual. Umbrella trials (in the same vein) like Lung-MAP assigned patients with the same tissue diagnosis (and often the same treatment approach, e.g., NSCLC) to different arms based on their mutation profiles [31]. The design of these trials allows for the characterization of specific heterogeneity within a single tumor type while supporting a platform-based approach for the rapid evaluation of multiple targeted therapies within a shared enrollment infrastructure. Many trials have incorporated adaptive features, such as Bayesian arm expansion and pre-specified interim analyses, which enable early termination for futility or success. Adaptive approaches bring an efficiency to resource utilization. Oxford's FOCUS4 trial in colorectal cancer is one of the first adaptive trials with genomic stratification and treatment arms that can be opened or closed based on interim data [32]. Similarly, Cambridge and UCL's TRACERx lung cancer study tracks clonal evolution over time while also incorporating longitudinal genomics to guide treatment adaptation [12]. These trials have established real-time sequencing, adaptive randomization, and interim futility analyses as key pillars of contemporary clinical trial design. Additionally, hybrid trial designs that incorporate additional layers of data, such as epigenetic profiles, immune landscapes, and spatial transcriptomics, are being employed to enable multi-modal patient 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 for the 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, linked to NHS records in Oxford and Cambridge, will enable the identification of matched patients as soon as actionable alterations are detected. Cohorts such as these benefit from

2025, Vol. 2, Article ID. 2025.0028
https://doi.org/10.69709/CAIC.2025.144976

integration using data standards like GA4GH (Global Alliance for Genomics and Health), enabling trial sites to communicate with each other. Profiling is also aided by the availability of liquid biopsy tests, which allow the detection of ctDNA mutations in real time during 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 using these methods of real-time stratification and treatment guidance [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 tumor biology. This is an example of 'real-time eligibility,' in which eligibility criteria are variable and depend on the tumor's molecular status over time.

3.3. Overcoming Barriers: Equity, Interpretation, and Scalability

However, challenges remain despite these relative technical successes. Limited access and underrepresentation of ethnic minority and rural populations can reduce generalizability [5]. This is exacerbated by issues with access to sequencing infrastructure and by the institutions conducting the trial research. These institutions, such as Harvard and MIT, are piloting mobile phlebotomy units and digital consent processes to decentralize the recruitment process and potentially eliminate geographic and socioeconomic impediments. The GenOMICC project in Oxford also represents an example of striving to include different ancestries to include severe cancer phenotypes and give examples of inclusive trial design. Furthermore, linkage with biobank initiatives (e.g., UK Biobank and All of Us Research Program) provides 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 in algorithmically enhanced, tumor-agnostic, and continuously adaptive master protocols that can account for the complexities of cancer biology and the ongoing 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 treatment response data. The DARPA-enabled Intelligent Trial Design program at MIT is developing self-optimizing systems to allocate patients across arms, leveraging omics data, patient clinical trajectories, and patient-reported outcomes [37]. These frameworks use reinforcement learning and Bayesian optimization to leverage prior-arm knowledge to inform the structure of subsequent trials in real time [38]. The contributions of Oxford and Cambridge to the PAN-COVID cancer study have also informed frameworks for real-world trial extensions, enabling adaptive modifications 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 helps overcome obstacles to establishing collaborations between institutions, which make 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 larger patient numbers. Apart from enhancing success rates and timelines for validating adaptive trials, these 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, enabling the realization of the full potential of precision oncology.

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

Precision medicine integrates molecular data, clinical context, and advanced analytics to tailor prevention, diagnosis, and treatment to individual patients. This branch of medicine is also revolutionizing diagnostic, treatment, and patient-tracking approaches by integrating with artificial intelligence (AI) to accurately diagnose chronic and complex diseases [39]. Cancer is biologically heterogeneous and remains a leading global cause of morbidity and mortality, with substantial variation in subtype, stage, and treatment response [40,41]. Moreover, this disease also varies significantly from person to person in terms of type, stage, and response to treatment [42]. Given this diversity, treatment cannot be standardized. Due to this diversity, treating such diseases has been difficult even for doctors ever since [40]. However, AI integration in such diseases has significantly eased several challenges [43]. Within the past 32 years, the mortality rate from such diseases has decreased by 33%. Advanced science, especially the evolution of precision medicine, has driven

improvements in these diseases. Advances in the study of such diseases have, in fact, led to more effective and more precise patient treatments through integration with AI. AI helps detect hidden patterns in images, estimate the progression of diseases, propose treatments, and determine whether a patient is eligible to enroll in clinical trials [44]. Indeed, in 2021, the FDA approved 71 AI-enabled devices. Moreover, more than 80% of such devices were used for cancer diagnosis. These devices were primarily utilized in radiology, pathology, and radiation oncology in cancers, especially solid cancers such as cancers in the breasts, lungs, and prostate [42,44]. Although these devices have been widely used to improve decision-making accuracy, their full implementation has been limited by potential algorithmic biases, inefficiencies in estimations, and increased burdens on healthcare systems [44].

4.1. AI in Patient Stratification and Trial Matching

One of the most influential applications of AI in precision oncology is the stratification of patients eligible for clinical trials based on their genomic, imaging, and clinical information. The complexity of cancer, both intra- and inter-tumoral, makes conventional patient stratification in clinical trials notoriously difficult, often leading to the exclusion of patients with unique profiles. AI helps overcome this hurdle by identifying hidden patterns in high-dimensional data, thereby improving patient stratification accuracy and inclusivity [45–47].

The DeepPatient model, trained on data from more than 700,000 electronic health records, demonstrated the ability of unsupervised learning to identify patient properties and disease patterns within 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 computation 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.

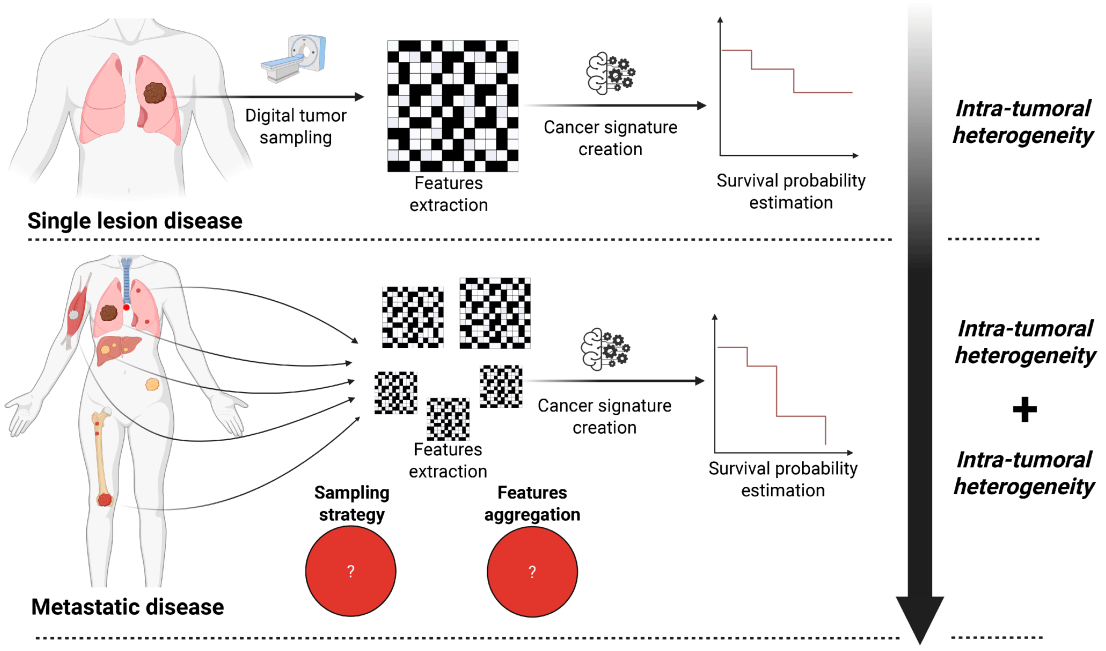

Figure 1: Comparison of radiomic analysis pipelines in single lesion and metastatic disease (reused from T. Henry et al., This is an open access article distributed under the terms of the Creative Commons CC BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided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

2025, Vol. 2, Article ID. 2025.0028
https://doi.org/10.69709/CAIC.2025.144976

4.2. AI-Powered Protocol Design and Adaptive Learning Systems

The inflexibility of conventional clinical trial methods limits their ability to adapt to evolving patient responses and new biomarker discoveries. By enabling the development of real-time, adaptive trial methods, AI overcomes this limitation. 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 stopping are among the trial flow optimization techniques employed by reinforcement learning, a subfield of machine learning in which computers learn through 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, and 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 compared to PFS1 (Von Hoff model).	RNA and DNA profiling help improve treatment matching.	Did not meet primary endpoint; requires large-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 are activated rapidly once a matching patient is found. Useful for rare genotypes.	Speeds up trial access for rare cases.	Site setup delays are still possible; early patient identification is not always feasible.

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

The current era of precision oncology also demands the integration of multi-modal data, including genomics, transcriptomics, proteomics, imaging, digital health data and patient-reported outcomes, to personalize both clinical trial design and treatment strategies. Figure 2 represents the paradigm of personalized medicine, in which treatment approaches are informed by genetic, environmental, 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 integrating multi-omic in-

formation with graph neural networks and deep learning to classify disease subtypes, 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) provide insights with greater accuracy than current clinical trial cohorts. AI-driven multimodal approaches also enable the use of real-world evidence (RWE) in regulatory settings to support the approval of new drugs and facilitate post-market surveillance [57]. However, certain limitations remain, including the need to standardize data, ensure AI model interpretability, preserve data, and enable data exchange across multiple institutions, which must be overcome to apply multimodal AI in clinical trials.

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) -plasma MATCH	-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 treatment responses using comprehensive tumor sequencing to identify predictive mutations.	Identify strong predictive biomarkers, understand drug mechanisms, and reduce trial size and cost.	Small sample size, limited data harmonization, and 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, and 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, and broaden 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, and increase quality of care.	Expensive to implement, digital literacy issues, less precise self-reporting, and rarely accepted by regulators.

2025, Vol. 2, Article ID. 2025.0028
https://doi.org/10.69709/CAIC.2025.144976

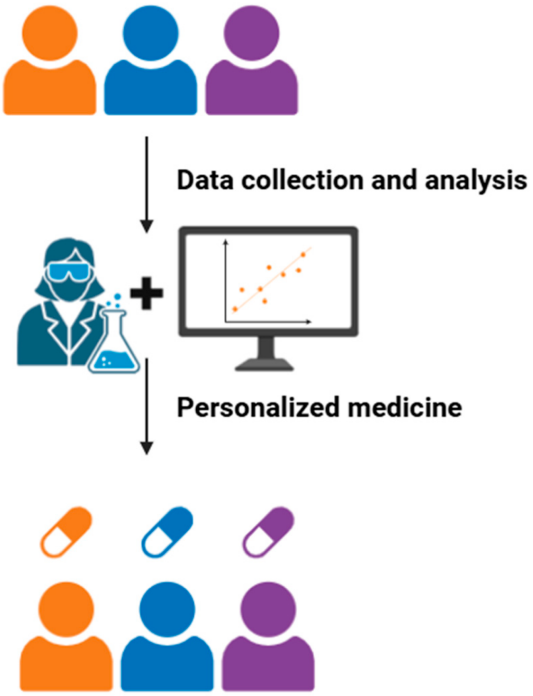

Figure 2: Personalized medicine framework integrating genetic, environmental, and lifestyle variables.

4.4. Ethical, Regulatory, and Operational Challenges

Although AI holds significant transformative potential for the design of precision oncology trials, its successful implementation depends on addressing key ethical, regulatory, and operational challenges. These challenges arise from the complexities of applying AI within healthcare systems and the inherent characteristics of the data.

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

Algorithmic bias ranks among the most pressing ethical issues, as AI models rely on homogeneous and/or incomplete data, leading to inequitable outcomes. For example, incomplete data on ethnic minorities 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 for understanding clinical perceptions and informed consent when AI models participate in trials and treatment. Emerging alternatives to AI models, termed Explainable AI (XAI) and continuous consent models, have been proposed to produce interpretable results while maintaining performance, 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 hand, actions taken by AI models might be clarified by specifying which change in input would cause an altered outcome in classification. Explanations must then be tracked in the clinical environment to ensure they meet trial thresholds for triage, fairness, and drift [60,61].

4.4.2. Regulatory Challenges: Lack of Standard Frameworks

There is an evolving landscape for 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 a guideline on the use of AI throughout 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 regulations and standards are established, the responsibility for ensuring compliance with AI in clinical trials rests with the sponsors and investigators.

4.4.3. 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 the 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 that is still largely absent in most current trial infrastructures. 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 convergence of artificial intelligence, multi-omics data, digital health platforms, and decentralized clinical trial infrastructure is transforming the future of clinical tri-

als. 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 change in clinical trials is DIGITAL TWINS. Digital twins are computational models of patients/cohorts that forecast disease progression, treatment response, and treatment risk, using data input parameters specific to each patient/cohort. 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 or storage of patient data in central databases. Thus, continuous learning through collaborations across multiple data points in trials conducted globally, while preserving 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-constrained settings.

5. Institutional Case Studies in Translational Precision Oncology

Even if precision oncology were a global affair, its history has been disproportionately shaped by certain institutions, such as Harvard University, the Massachusetts Institute of Technology (MIT), Oxford University, and Cambridge University. These serve as hubs of innovation, not only by advancing our understanding of the genome, but also by developing clinical, computational, and ethical frameworks that are redefining the practice of cancer science.

5.1. AI-Enabled Clinical Stratification and Real-World Evidence Integration

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 benchmarks for mutation prevalence 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 for the biomarker-first stratification frameworks now used worldwide. Further, the creation of MatchMiner, an open-source, artificial intelligence-enabled clinical trial matching platform, is driving trial enrolment and increasing equitable 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 affiliated with Harvard are increasingly utilizing multi-omic dashboards powered by explainable AI. This allows oncologists to use transcriptomics, radiomics, and proteomics in additive ways to inform real-time treatment decisions.

5.2. Computational Engineering and Digital Twin Technologies in Oncology

MIT's unique potential lies in its integrated approach to engineering, AI, and biomedical science. For example, the Koch Institute has contributed to numerous advancements, including tumor-targeted nanoparticles, programmable drug delivery, and synthetic biology diagnostics. The institute is also at the preclinical stage of developing technologies for early diagnosis and trial stratification, such as CRISPR-Cas sensors for detecting circulating tumor DNA [64].

The Jameel Clinic (referred to as J-Clinic) collaborates with graph neural networks (GNNs) and deep reinforcement learning to develop models of virtual patients, design adaptive trials, and predict adverse event risks before patients receive treatment [12,23]. These technologies also enable the optimization of dosing schedules, reduce the risk of participant dropout after randomization, and predict potentially synergistic drug interactions. With support from the FDA Oncology Center of Excellence, MIT's digital twin program enables researchers 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 involves digital pathology, patient-reported outcomes, and the calibration of digital twins using time-series imaging, such as [65].

2025, Vol. 2, Article ID. 2025.0028
https://doi.org/10.69709/CAIC.2025.144976

5.3. Genomic Infrastructure, Federated Data Governance, and Ethical AI

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 investigators contributed to the 100,000 Genomes Project, where it has innovated somatic variant detection, structural variation, and clinical-grade reporting pipelines [22, 68]. 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 conduct translational research and advance it to regulatory-grade standards. 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. Multimodal Biomarker Discovery and Clinical Implementation Science

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 used liquid biopsy endpoints to create adaptive cohorts in the neoadjuvant and postoperative contexts [69]. The RadNet program develops and evaluates innovations in radiogenomics, spatial transcriptomics, and immune landscape characterization to identify predictors of radioresistance. Cambridge is also validating the use of exosomal RNA and fragmentomics as surrogate trial endpoints through the CAPTURE and SIGNATURE studies, technologies that are being adopted by pan-European biomarker consortia [70]. Cambridge is also at the forefront of AI-powered clinical platforms. Its ongoing work on automated tumor boards, which integrates imaging, histopathology, and genomics through multimodal neural networks, is currently under evaluation by ESMO and the UK NHS Cancer Alliances for potential national deployment [30]. Beyond its academic contributions, Cambridge's active partnerships with AstraZeneca, Illumina, and GRAIL on the Biomedical Campus are helping bridge the gap between discovery and implementation science, enabling carefully phased first-in-human trials.

5.5. Synthesis

Each institution contributes uniquely to the shared mission of scalable, equitable, and ethically responsible translational oncology: Harvard with its AI-powered clinical systems, MIT with programmable AI and digital twins, Oxford with federated data governance, and Cambridge with multi-modal discovery and implementation of biomarkers. Together, these institutions exemplify a 'living' blueprint for how academic, clinical, and regulatory entities can co-evolve to deliver radically personalized cancer care at an unprecedented scale.

6. Future Perspectives

The evolution of precision oncology is no longer aspirational; it is 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 advance synchronously across three converging dimensions: 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 updates, employ synthetic control arms instead of traditional comparators, which offer both increased efficiency and greater ethical alignment [71]. With the ever-increasing application of spatial omics, quantum-inspired algorithms and long-read sequencing will offer greater insight into tumor microenvironments, immune niches and rare subclonal populations [72]. Interoperability will depend on cloud platforms and the use 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 that combines 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 harmonized 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 synchronize 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

The under-representation of minority and low-resource populations risks generating biased predictions and results in poor model 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 the 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 [62,81].

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, in developing countries, and via virtual platforms. The determination of whether precision oncology is a universal right or only a distinct privilege will hinge on integrating science, data, and ethics, bolstered by inclusive institutions, policy foresight, and civic trust.

7. Discussion and Conclusions

In conclusion, this paper has presented the revolutionary changes brought about by the convergence of genetic innovation, adaptive trials, and institutional developments in the field of precision oncology. At its core, this revolution applies AI, real-world evidence, and HTS to transform biomarker-based approaches from rigid procedural routines in precision oncology. Case studies involving Harvard, MIT, Oxford, and Cambridge universities represent the scientific acuity and translational foresight in integrating 'omics' data in clinical routines. However, there remain significant hindrances on the path. Rather, the current models in precision oncology, in terms of generalizability, continue to face inequities in patient enrollment, variant misclassification, and regulatory discordance. These points, specifically, serve as current limitations. Thus, the need arises to incorporate flexibility into future structuring, given the strongly important ethical, regulatory, and policy considerations. There may be a benefit to 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 leading institutions in the world, although one surmises

personal conviction in assuming major global promise in resolution. This paper presents the role of innervation in inclusive design, ethical necessity, and infrastructural necessity in concert with scientific advancement. Rather, with proper technologies, such as AI, liquid biopsy, and single-cell genomics, will be realized. Broadly, maintaining integrity to meet global regulatory standards, while addressing public trust, clinical need, and effective structuring, will play a major role in shaping future cancer treatment options. Rather, with the convergence of science in treatment, precision-to-global unanimity represents the major promise in attaining global recognition in mainstream treatment options for 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., Visualization 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.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

The study did not receive any external funding and was conducted using only institutional resources.

Acknowledgments

Declared None.

AI Declaration

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

[2] The Cancer Genome Atlas Network, "Comprehensive molecular portraits of human breast tumours," *Nature*, vol. 490, no. 7418, pp. 61–70, 2012. [CrossRef]

[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, 2017. [CrossRef]

[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," *Ann. Oncol.*, vol. 35, no. 7, pp. 588–606, 2024. [CrossRef]

[5] D. E. Spratt et al., "Racial/ethnic disparities in genomic sequencing," *JAMA Oncol.*, vol. 2, no. 8, pp. 1070–1074, 2016. [CrossRef]

[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, 2019, Art. no. 11. [CrossRef]

[7] M. Gerlinger et al., "Intratumor heterogeneity and branched evolution revealed by multiregion sequencing," *N. Engl. J. Med.*, vol. 366, no. 10, pp. 883–892, 2012. [CrossRef]

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

[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, 2017. [CrossRef]

[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, 2016. [CrossRef]

[11] U.S. Food and Drug Administration, Center for Drug Evaluation and Research (CDER), "Adaptive designs for clinical trials of drugs and biologics guidance for industry," 2019. Accessed: Jul. 28, 2025. [Online]. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics-guidance-industry

[12] M. Jamal-Hanjani et al., "Tracking the evolution of non–small-cell lung cancer," *N. Engl. J. Med.*, vol. 376, no. 22, pp. 2109–2121, 2017. [CrossRef]

[13] Z. Zhang et al., "Genetic analyses support the contribution of mRNA N6-Methyladenosine (m6A) modification to human disease heritability," *Nat. Genet.*, vol. 52, no. 9, pp. 939–949, 2020. [CrossRef]

[14] F. Iorio et al., "A landscape of pharmacogenomic interactions in cancer," *Cell*, vol. 166, no. 3, pp. 740–754, 2016. [CrossRef]

[15] J. Barretina et al., "The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity," *Nature*, vol. 483, no. 7391, pp. 603–607, 2012. [CrossRef]

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

[17] A. Roth et al., "PyClone: Statistical inference of clonal population structure in cancer," *Nat. Methods*, vol. 11, no. 4, pp. 396–398, 2014. [CrossRef]

[18] O. Shalem et al., "Genome-scale CRISPR-Cas9 knockout screening in human cells," *Science*, vol. 343, no. 6166, pp. 84–87, 2014. [CrossRef]

[19] J. M. Dempster et al., "Agreement between two large pan-cancer CRISPR-Cas9 gene dependency data sets," *Nat. Commun.*, vol. 10, no. 1, 2019, Art. no. 5817. [CrossRef]

[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, 2016. [CrossRef]

[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, 2018. [CrossRef]

[22] C. Turnbull et al., "The 100 000 genomes project: Bringing whole genome sequencing to the NHS," *BMJ*, vol. 361, 2018, Art. no. k1687. [CrossRef]

[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, 2020. [CrossRef]

[24] H. Gao et al., "The landscape of tolerated genetic variation in humans and primates," *Science*, vol. 380, no. 6648, 2023, Art. no. eabn8153. [CrossRef]

[25] J. Cheng et al., "Accurate proteome-wide missense variant effect prediction with alphaMissense," *Science*, vol. 381, no. 6664, 2023, Art. no. eadg7492. [CrossRef]

[26] G. M. Findlay et al., "Accurate classification of BRCA1 variants with saturation genome editing," *Nature*, vol. 562, no. 7726, pp. 217–222, 2018. [CrossRef]

[27] T. M. Gierahn et al., "Seq-well: Portable, low-cost RNA sequencing of single cells at high throughput," *Nat. Methods*, vol. 14, no. 4, pp. 395–398, 2017. [CrossRef]

[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, 2020. [CrossRef]

[29] J. E. Dahlman et al., "Barcoded nanoparticles for high-throughput in vivo discovery of targeted therapeutics," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 114, no. 8, pp. 2060–2065, 2017. [CrossRef]

[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," *Ann. Oncol.*, vol. 31, no. 11, pp. 1491–1505, 2020. [CrossRef]

[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)," *J. Thoracic Oncol.*, vol. 15, no. 10, pp. 1636–1646, 2020. [CrossRef]

[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," *Lancet*, vol. 390, no. 10114, pp. 2769–2778, 2017. [CrossRef]

[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, 2025. [CrossRef]

[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, pp. 155–166, 2021. [CrossRef]

[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, 2015. [CrossRef]

[36] L. Horn et al., "First-line atezolizumab plus chemotherapy in extensive-stage small-cell lung cancer," *N. Engl. J. Med.*, vol. 379, no. 23, pp. 2220–2229, 2018. [CrossRef]

[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, 2016, Art. no. 346ra92. [CrossRef]

[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, 2009. [CrossRef]

[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, 2021. [CrossRef]

[40] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011. [CrossRef]

[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, 2022, Art. no. FSO787. [CrossRef]

[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, 2022. [CrossRef]

[43] R. Hamamoto et al., "Application of artificial intelligence technology in oncology: Towards the estab-

lishment of precision medicine," *Cancers*, vol. 12, no. 12, 2020, Art. no. 3532. [CrossRef]

[44] L. Kolla and R. B. Parikh, "Uses and limitations of artificial intelligence for oncology," *Cancer*, vol. 130, no. 12, pp. 2101–2107, 2024. [CrossRef]

[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, 2020. [CrossRef]

[46] T. Miyauchi, T. Yaguchi, and Y. Kawakami, "Inter-patient and intra-tumor heterogeneity in the sensitivity to tumor-targeted immunity in colorectal cancer," *Jap. J. Clin. Immunol.*, vol. 40, no. 1, pp. 54–59, 2017. [CrossRef]

[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, 2022, Art. no. 20650. [CrossRef]

[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, 2016, Art. no. 26094. [CrossRef]

[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, 2019. [CrossRef]

[50] B. Koçak, E. Ş. Durmaz, E. Ateş, and Ö. Kılıçkesmez, "Radiomics with artificial intelligence: A practical guide for beginners," *Diagn. Interv. Radiol.*, vol. 25, no. 6, pp. 485–495, 2019. [CrossRef]

[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, 2020. [CrossRef]

[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, 2021. [CrossRef]

[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, pp. 303–310, 2019. [CrossRef]

[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, 2022. [CrossRef]

[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, 2023, Art. no. 29957. [CrossRef]

[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, 2021, Art. no. 5275. [CrossRef]

[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, 2024, Art. no. e13903. [CrossRef]

[58] J. L. Cross, M. A. Choma, and J. A. Onofrey, "Bias in medical AI: Implications for clinical decision-making," *PLOS Digit. Health*, vol. 3, no. 11, 2024, Art. no. e0000651. [CrossRef]

[59] Z. Sadeghi et al., "A review of explainable artificial intelligence in healthcare," *Comput. Electr. Eng.*, vol. 118, 2024, Art. no. 109370. [CrossRef]

[60] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should i trust you?' explaining the predictions of any classifier," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, Aug. 13–17 2016, pp. 1135–1144. [CrossRef]

[61] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.*, pp. 4766–4775, 2017. [View Online]

[62] U.S. Food and Drug Administration, "Artificial intelligence in software as a medical device," Accessed: Jul. 28, 2025. [Online]. Available: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device

[63] European Medicines Agency, "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*, vol. 594, no. 7862, pp. 265–270, 2021. [CrossRef]

[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 Digit. Oncol.*, vol. 5, 2024, Art. no. 100056. [CrossRef]

[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. [CrossRef]

[67] R. E. Sherman et al., "Real-world evidence - what is it and what can it tell us?" *N. Engl. J. Med.*, vol. 375, no. 23, pp. 2293–2297, 2016. [CrossRef]

[68] E. Pairo-Castineira et al., "Genetic mechanisms of critical illness in COVID-19," *Nature*, vol. 591, no. 7848, pp. 92–98, 2020. [CrossRef]

[69] Cancer Research UK, "Tracking cancer evolution: The TRACERx lung study," 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] Cancer Research UK, "RadNet-our radiation research network," 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, 2023. [CrossRef]

[72] V. Marx, "Method of the year: Spatially resolved transcriptomics," *Nat. Methods*, vol. 18, no. 1, pp. 9–14, 2021. [CrossRef]

[73] J. Qin, Y. Chen, Y. Peng, Z. Wang, and L. Wang, "Freeze the backbones: A parameter-efficient contrastive approach to robust medical vision-language pre-training," in *Proc. ICASSP, IEEE Int. Conf. Acoust. Speech Signal-Process.-Proc.*, Barcelona, Spain, May 4–8, 2024, pp. 1686–1690. [CrossRef]

[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, pp. 14751–14773, 2023. [View Online]

[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, 2018. [CrossRef]

[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, pp. 1–23, 2022. [CrossRef]

[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, pp. 347–368, 2023. [CrossRef]

[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, 2021. [CrossRef]

[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, 2019. [CrossRef]

[80] S. Taylor, "'Global health': Meaning what?" *BMJ Glob. Health*, vol. 3, no. 2, 2018, Art. no. e000843. [CrossRef]

[81] European Commission, "AI act | shaping europe's digital future," Accessed: Sep. 30, 2025. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai