

PUBLIC SUBMISSION

Received: April 30, 2025
Tracking No. ma3-nb4l-ss8t
Comments Due: May 28,
2025 **Submission Type:** API

Docket: NSF-2025-OGC-0001
NITRD_FRDOC_0001

Comment On: NSF-2025-OGC-0001-0001
Request for Information: Development of a 2025 National Artificial Intelligence Research and Development Strategic Plan

Document: NSF-2025-OGC-0001-DRAFT-0016
Comment on FR Doc # 2025-07332

Submitter Information

Organization: Massive Bio, Inc.

General Comment

On behalf of Massive Bio, I am honored to submit our response to the 2025 National AI R&D Strategic Plan RFI. As Co-Founder and Chief Medical Officer, I emphasized how AI—when properly integrated into clinical research and care delivery—can revolutionize cancer care and health equity.

We strongly support the Administration's vision outlined in Executive Order 14179. Our response highlights strategic areas where federal investment is critical:

- Deployment of digitally enabled AI prescreening hubs to manage data drift and streamline real-time clinical trial matching.
- Making clinical trials a flagship use case for national EMR interoperability, privacy-centric consent, and patient-facing tools.
- Foundational AI for modeling complex biological systems and accelerating personalized medicine.
- AI to improve equity and efficiency in clinical trials, including intelligent trial design and predictive analytics.
- Explainable, fair, and robust AI to unlock real-world evidence and support trustworthy deployment.
- Secure data infrastructure with federated learning, synthetic data, and standardized APIs.

We also propose mechanisms like national data collaboratives, AI grand challenges, regulatory sandboxes, and workforce development. Massive Bio is committed to advancing this national priority in partnership with OSTP, NITRD, and other stakeholders.

Please find our full comment attached for your review.

Warm regards,
Dr. Arturo Loaiza-Bonilla
Co-Founder & Chief Medical Officer
Massive Bio, Inc.

Attachments

AI and Digital Prescreening hubs 2025

Dr. Loaiza-Bonilla - ACS_CoC_Reducing Barriers to Clinical Trial Enrollment through AI_Massive Bio_AI 4.25.2025

AIp2500005

AI NSF R_D Action Plan_Massive Bio_4.2025

CLOSING THE GAPS IN ONCOLOGY:



HOW AI AND DIGITAL PLATFORMS ARE TRANSFORMING PATIENT ADHERENCE AND REAL-TIME DECISION SUPPORT

By Arturo Loaiza-Bonilla,
MD, MSEd

Precision oncology has brought about a paradigm shift in cancer care by harnessing molecular profiling and biomarker-driven therapies.

Through the identification of genetic or protein-based alterations, these targeted treatments are tailored to individual patient profiles rather than relying solely on traditional chemotherapy or radiation.

While this approach has shown improved outcomes in specific cancer subtypes, its real-world application faces persistent hurdles — most notably delays in prior authorization,

inconsistent access to next-generation sequencing (NGS) and inefficient data-sharing infrastructures.^{1,2}

An additional layer of complexity arises from the challenge of ensuring patient adherence to therapies, particularly oral oncolytics, which patients typically self-administer outside of the clinical setting.²



Arturo Loaiza-Bonilla

Incomplete data on side-effect management, financial hurdles (such as co-payments or insurance denials) and logistical issues (like long travel times to infusion centers) often compound these obstacles. Collectively, these barriers compromise the timely delivery of precision therapies and curtail the full realization of personalized cancer medicine.

Fortunately, technological solutions powered by artificial intelligence (AI) are emerging to streamline these processes. AI-driven clinical decision support systems (CDSSs) and digital patient engagement platforms promise to close critical gaps in data accessibility, patient adherence and treatment optimization. Early evidence suggests that these solutions can significantly reduce time-to-therapy, minimize discontinuations, and improve patient satisfaction.^{3,4}

By integrating real-time data from various sources — including electronic health records (EHRs), diagnostic laboratories and telemedicine portals — AI solutions are well-positioned to transform reactive oncology care into a proactive, patient-centered model.^{1,5}

CONTINUED ON NEXT PAGE

CLOSING THE GAPS

CONTINUED FROM PREVIOUS PAGE

CHALLENGES IN PRECISION ONCOLOGY: THE INFORMATION BOTTLENECK

Despite significant advances, precision oncology often operates in a fragmented healthcare environment. As the complexity of treatment decisions has grown, clinicians increasingly rely on comprehensive, up-to-date data to choose the best therapeutic pathway. Key bottlenecks include:

▲ Fragmented Health Records:

In many healthcare systems, patient data remains siloed in disparate EHR platforms and testing laboratories, impeding efficient care coordination. This fragmentation can lead to duplicated diagnostic tests, delayed biomarker results and incomplete documentation of adverse events — all of which delay optimal treatment starts.^{1,5}

▲ Underutilization of NGS Testing:

While clinical guidelines often recommend biomarker testing to identify actionable targets, utilization of these tests remains uneven. Disparities in access to testing resources, especially for patients in rural or underserved communities, create a gap where up to 30% to 40% of potentially eligible patients are never tested for certain biomarkers.^{4,5}

▲ Adherence and Non-Completion:

Even if a precision therapy is identified, adherence to oral oncolytics and targeted agents remains an underappreciated challenge. Patients may discontinue treatment due to side effects, insurance coverage lapses or a lack of real-time support at home. Some studies estimate that more than 20% of patients prescribed oral anticancer medications exhibit nonadherence, leading to suboptimal outcomes.^{2,3}

▲ Dependence on Retrospective Claims Data:

Many analyses guiding therapeutic decision-making and commercial efforts are based on retrospective insurance claims. While claims data can provide high-level insights, they are rarely in real time and often miss critical details like patient-reported symptoms, molecular subtyping

and social determinants of health. This temporal lag hinders proactive interventions that could improve therapeutic decision-making and reduce attrition.^{1,6}

Collectively, these challenges highlight the urgent need for integrated digital frameworks capable of delivering real-time, patient-specific recommendations and alerts.

AI-Driven Frameworks for Optimizing Oncology Care

REAL-TIME CLINICAL DECISION SUPPORT

AI-based CDSS platforms leverage algorithms that analyze patient-level data (including medical history, biomarker profiles and treatment responses) to provide evidence-based recommendations. By integrating data from EHRs, pathology labs and imaging facilities, these platforms can:

▲ Identify Under-Genotyped Populations:

When NGS testing rates lag behind guideline recommendations, AI can proactively flag and prioritize patients who have not received comprehensive molecular profiling.⁴

▲ **Highlight Therapy Eligibility:** Predictive analytics can alert oncologists to potential matches for targeted therapies or clinical trials, even when these options have not been top-of-mind in a busy clinical setting.⁵

▲ **Optimize Drug Utilization:** Advanced platforms can integrate real-world biomarker trends and prescribing patterns, ensuring that patients who qualify for emerging therapies are identified quickly.⁶

An example of such a tool is Massive Bio's Drug Utilization Optimizer (DUO), which tracks drug usage and biomarkers in real time to reduce bottlenecks between diagnostic results and therapy selection.⁹

By highlighting immediate opportunities for targeted treatment, DUO serves as a bridge between molecular diagnostics and clinical implementation, enhancing the likelihood that patients receive the most effective therapy at the right moment.

Additionally, AI-based image analysis

for biomarkers such as PD-L1 expression can reduce interobserver variability, expediting decisions on immunotherapy eligibility and bridging the gap between pathology workflows and oncology clinics.^{1,5}

PATIENT ENGAGEMENT AND ADHERENCE

Even in ideal scenarios where a patient is matched to the most suitable targeted therapy, sustained adherence is essential for meaningful clinical benefit. AI-powered digital tools and chatbots address this by:

▲ **Providing Proactive Reminders:** Through smartphone notifications, patients can receive personalized prompts to take their medications, report side effects or schedule routine lab tests. These reminders can be adapted based on each patient's dosing schedule and risk factors for noncompliance.²

▲ **Side-Effect Management:** In-app triage algorithms can guide patients through common adverse events such as nausea, fatigue or rash, and connect them in real time with supportive care or oncology nurse navigators.³

▲ **Financial Support:** Many patients are unaware of assistance programs or may struggle with navigating copay systems. AI-driven chatbots can direct patients to relevant resources, expediting financial approvals and mitigating therapy abandonment due to cost.^{2,7}

▲ **Predictive Dropout Alerts:** By analyzing usage patterns, AI platforms can detect early signals of potential nonadherence — such as missed doses or unaddressed side effects — and alert care teams before problems lead to therapy discontinuation.³

Massive Bio's AskFiona AI and DrArturo AI chatbots extend this approach by providing real-time, patient-centered support for individuals receiving targeted therapies or participating in clinical trials.⁹ These chatbots incorporate educational materials, medication reminders and direct referral pathways to biomarker testing and support services, aiming to streamline patient engagement throughout the

CONTINUED ON NEXT PAGE

CLOSING THE GAPS

CONTINUED FROM PREVIOUS PAGE

treatment cycle. Early evidence suggests that integrating AI chatbots into the routine oncology workflow can boost adherence and improve patient satisfaction — outcomes that directly influence both short- and long-term prognosis.^{2,3,9}

CLOSING THE LAST-MILE GAP

Even when a patient's therapy selection and supportive care plan are in place, logistical factors can hinder timely treatment initiation. AI-driven prescreening hubs aim to track patient progress in real time, considering both clinical trial enrollment opportunities and standard-of-care options.^{1,4} These hubs:

▲ **Monitor Disease Progression:** Ongoing data collection from wearable devices, lab results and imaging can reveal subtle shifts in disease status that signal the need for therapy adjustments.^{3,4}

▲ **Coordinate Across Care Teams:** Automated dashboards can systematically notify medical, surgical and radiation oncologists — and even commercial or manufacturer support teams — when a patient's treatment plan needs urgent attention.^{5,6}

▲ **Streamline Testing and Referrals:** For patients in remote areas or smaller community centers, AI can identify the nearest biomarker testing facilities and connect them with telemedicine consultations. This reduces travel burdens and ensures equitable access to precision therapies.^{4,7}

By bridging these final logistical hurdles, comprehensive AI platforms — including DUO and integrated chatbots — help ensure that patients do not “fall through the cracks” during transitions between diagnostic tests, insurance approvals and therapy initiation.

REGULATORY AND PRIVACY CONSIDERATIONS

As AI rapidly integrates into oncology, maintaining robust regulatory and ethical standards is paramount:

▲ **Data Privacy and Security:** Platforms must comply with HIPAA, GDPR and other

Ultimately, these innovations signal a shift from a static model of oncology care toward a continuously learning healthcare system. As AI refines the precision of cancer treatment, the role of clinicians evolves to focus more on nuanced clinical judgments, empathetic communication and shared decision-making with patients.

regional regulations to protect sensitive patient data.^{2,7} Encrypted data storage, strict access controls and rigorous de-identification measures are essential.

▲ **Clinical Validation and FDA Oversight:** AI-driven decision support systems that directly influence clinical care may require regulatory review or U.S. Food and Drug Administration (FDA) clearance, depending on their level of autonomy and impact on patient outcomes.^{5,6}

▲ **Ethical Frameworks:** Transparent algorithm design is vital to build trust among patients and providers. Emphasizing patient-consented data sharing ensures that individuals retain control over how their health information is used.^{7,8}

▲ **Bias and Equity:** AI algorithms can inadvertently exacerbate existing health disparities if trained on unrepresentative datasets. Continuous audit and review of AI tools are essential to mitigate bias in patient identification and therapy recommendations.^{3,7}

As these frameworks are further

refined, real-world evidence (RWE) from post-market studies and pragmatic trials will likely shape the next generation of AI regulatory guidelines.

FUTURE DIRECTIONS

The trajectory of AI in precision oncology is poised for significant expansion over the coming decade:

▲ **Predictive Risk Stratification:** By integrating multiomics data — such as genomics, proteomics and metabolomics — AI can predict disease risk and response profiles more precisely, leading to earlier interventions and personalized surveillance strategies.⁴

▲ **Synthetic Data Generation (Digital Twins):** Utilizing virtual patient models, researchers can model hypothetical treatment courses, optimizing clinical trial design and accelerating approval for promising drugs.^{3,8}

▲ **Expanded Chatbot and Telehealth Functionalities:** Future chatbots will likely incorporate psychosocial screening, financial navigation, and integration with real-time biomarker test scheduling. This “omni-channel” approach makes advanced oncology support accessible to patients in both urban and remote settings.^{2,7}

▲ **AI-Guided Adaptive Trials:** As we move toward master protocols and basket trials, AI can dynamically update eligibility criteria based on accumulating data, improving trial efficiency and reducing patient burden.^{1,5}

Ultimately, these innovations signal a shift from a static model of oncology care toward a continuously learning healthcare system. As AI refines the precision of cancer treatment, the role of clinicians evolves to focus more on nuanced clinical judgments, empathetic communication and shared decision-making with patients.

CONCLUSION

AI-driven clinical decision support and digital engagement tools offer transformative potential in today's oncology landscape. By delivering real-time alerts on therapy eligibility, simplifying prior authorization workflows and promoting

CONTINUED ON NEXT PAGE

CLOSING THE GAPS

CONTINUED FROM PREVIOUS PAGE

patient adherence through personalized digital platforms, these solutions address some of the most critical bottlenecks in precision oncology.¹⁻⁸

Crucially, AI does not aim to replace human expertise but rather to augment and streamline it, ensuring that medical teams can focus on high-level clinical reasoning and patient rapport. The benefits of this paradigm shift extend across the oncology ecosystem:

▲ **Patients** receive timely, targeted interventions supported by continuous engagement.

▲ **Clinicians** gain actionable insights at the point of care without sifting through disparate data systems.

▲ **Industry stakeholders** — from drug manufacturers to payers — benefit from enhanced treatment fidelity, higher-quality real-world evidence and reduced waste.

Most importantly, patient outcomes improve when therapy selection is refined by robust data, timely intervention and proactive adherence measures.

Realizing this vision at scale will

require ongoing collaboration among technology providers, healthcare institutions, regulatory bodies and patient advocacy groups. As precision oncology continues to evolve, efforts to incorporate next-generation AI platforms must be matched by stringent ethical frameworks, data privacy safeguards and equity-focused deployment.

By meeting these challenges, we can ensure that AI-driven precision oncology becomes a widely accessible standard, delivering on its promise of improving both survival and quality of life for cancer patients worldwide.

▲ **Arturo Loaiza-Bonilla, MD, MSc,** is co-founder and Chief Medical Officer of Massive Bio, New York, New York, and Network Chief of Hematology and Oncology, St. Luke's University Health Network, Easton, Pennsylvania.

REFERENCES

1. Artificial Intelligence in clinical decision support systems for oncology. PubMed, <https://pubmed.ncbi.nlm.nih.gov/36619220/>.
2. An AI-driven solution to supporting oncology oral therapy adherence. Memora Health, <https://www.memorahealth.com/resource/an-ai-driven-solution-to-supporting-oncology-oral-therapy-adherence>.

3. AI-Augmented Clinical Decision Support in a Patient-Centric Framework. Liebert Pub, <https://www.liebertpub.com/dol/10.1089/alpo.2023.0001>.

4. Top 10 AI Contributions to Cancer Research and Treatment. Augnito, <https://augnito.ai/resources/10-ai-led-contributions-to-cancer-treatment/>.

5. A clinical decision support system for AI-assisted decision-making in oncology. Nature, <https://www.nature.com/articles/s41598-023-32032-6>.

6. Effect of an Artificial Intelligence Clinical Decision Support System on Treatment Decisions and NCCN Guideline Adherence for Patients With Breast Cancer. ASCO Publications, <https://ascopubs.org/dol/10.1200/JCO.20.00018>.

7. Decision Support Systems in Oncology. ASCO Publications, <https://ascopubs.org/dol/10.1200/JCO.18.00001>.

8. Use and accuracy of decision support systems using artificial intelligence in oncology: A systematic review. Frontiers in Oncology, <https://www.frontiersin.org/articles/10.3389/fonc.2023.1224347/full>.

9. DUO: Drug Utilization Optimizer and Chatbots by Massive Bio, <https://www.youtube.com/watch?v=PZnKXboWo8E> (Accessed on February 18, 2025).





  minnesota cancer alliance

The Minnesota COC Network Presents

Reducing Barriers to Clinic Trial Enrollment through AI

Friday, April 25, 9am - 10:30 am CST


Arturo Loaiza-Bonilla, MD, MEd


Pia Banerjee, PhD



Reducing Barriers to Clinical Trial Enrollment through AI Solving the Paradox: The Future of Clinical Trial Matching

Arturo Loaiza-Bonilla, MD, MEd, FACP

Solving the Paradox: The Future of Clinical Trial Matching



Arturo Loaiza-Bonilla, MD, MEd, FACP

- Chief of Hematology and Oncology, St. Luke's University Health Network
- Co-Founder, Massive Bio
- Associate Professor at Lewis Katz School of Medicine at Temple University
- Past President, Pennsylvania Society of Oncology and Hematology – ASCO Affiliate



Disclosures

- **THIS PRESENTATION WILL BE RECORDED FOR SCIENTIFIC PURPOSES**
- **Consulting: Verily (Google/Alphabet), MedScape/WebMD, PSI CRO, Astra Zeneca, Ipsen, BrightInsight, Science37**
- **Speaker's Bureau: Guardant**
- **Stock: Massive Bio, BioLineRx**

“This presentation reflects my own personal view on the issues discussed and does not reflect the view of my employer, St. Luke’s University Health Network and its affiliates.”

Oncology has experienced two of the greatest technological evolutions: molecular "omics" (genomics, proteomics, epigenomics) and "big data"

- A couple of decades ago, cancer was diagnosed using a combination of X-ray imaging and histopathology tests. In contrast, molecular tests can now report on changes in hundreds of genes and proteins to diagnose and determine the prognosis and treatment of cancer in an individual.
- In fact, these advances are extending survival and improving the quality of life of hundreds of thousands of patients, yet healthcare professionals face new challenges associated with the implementation of precision medicine,
- Growth of medical knowledge is exponential
- Constant specialization is required to provide highly individualized cancer care.
- This debate is not unique to highly developed countries. Providing comprehensive, state-of-the-art cancer care to millions of patients remains a significant challenge, particularly for suburban and rural populations.
- Biomedical data is heterogeneous and difficult to classify (e.g. high dimensionality, time dependence, parity, irregularity) for Artificial Intelligence applications.



Biomarker Technologies

Immunohistochemistry (IHC)

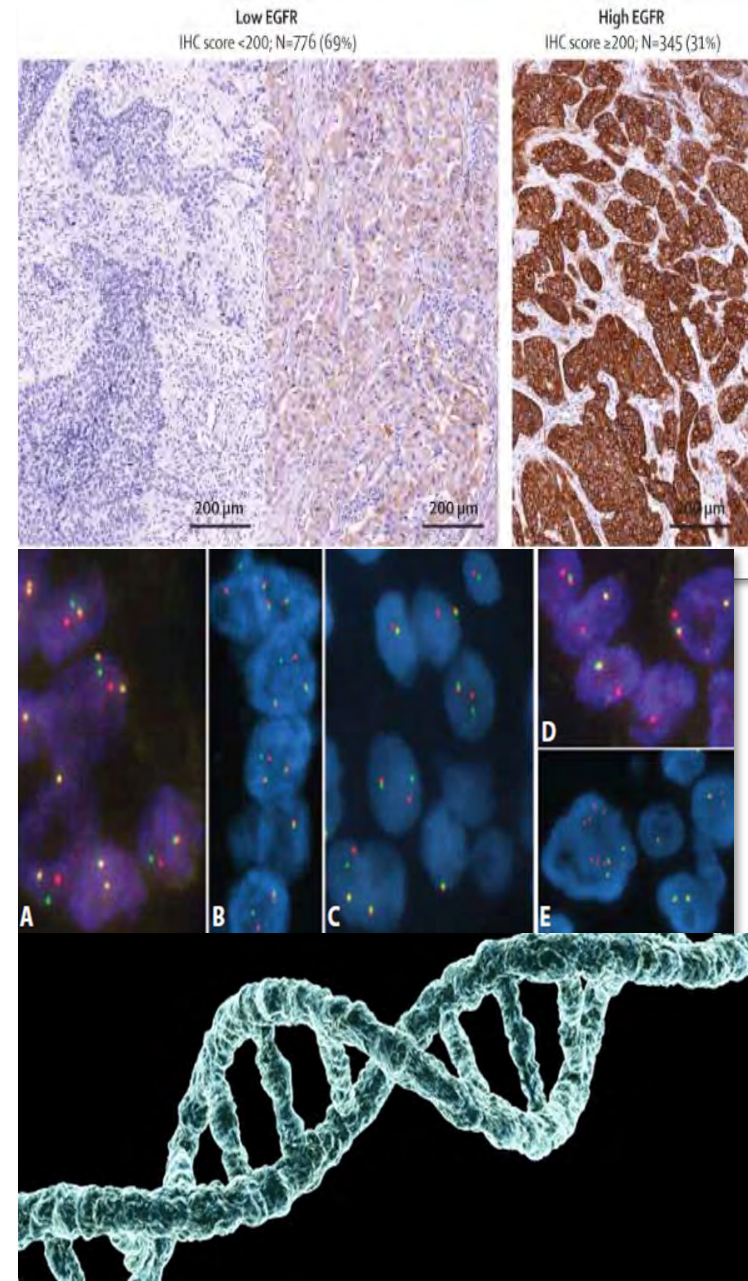
- Biomarker technique used to detect level of protein expression

Fluorescence In Situ Hybridization (FISH)

- Biomarker technique used to detect alterations in DNA (single-gene)

Next-Generation Sequencing (NGS)

- High-throughput biomarker technique used to detect alterations in DNA (multi-gene) and to construct a comprehensive genomic profile



Illumina Aims to Push Genetics Beyond the Lab With \$200 Genome

- New machine pushes company toward goal of \$100 genome sequence
- Could be boon for gene-targeting drugs, Regeneron exec says

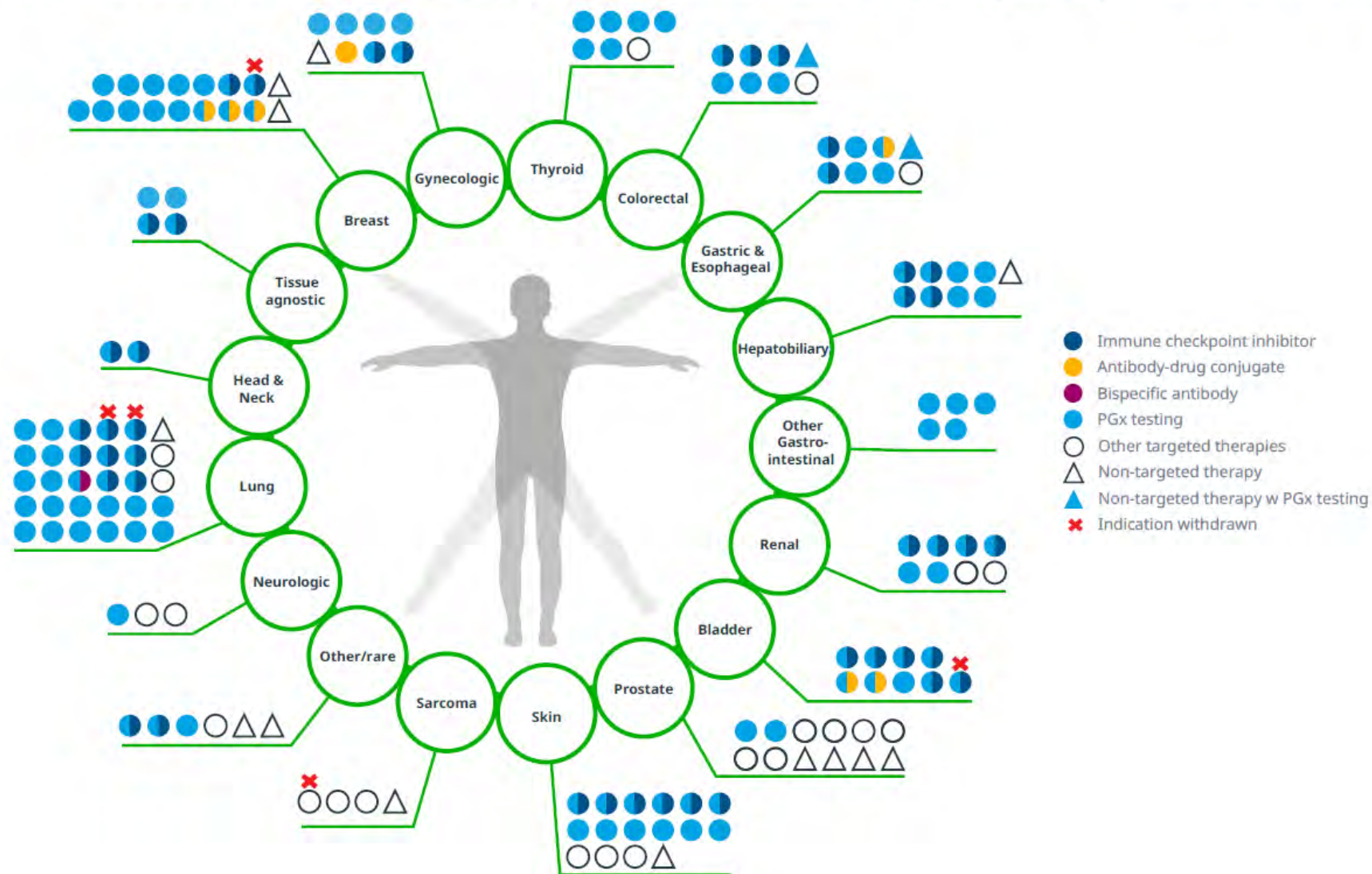


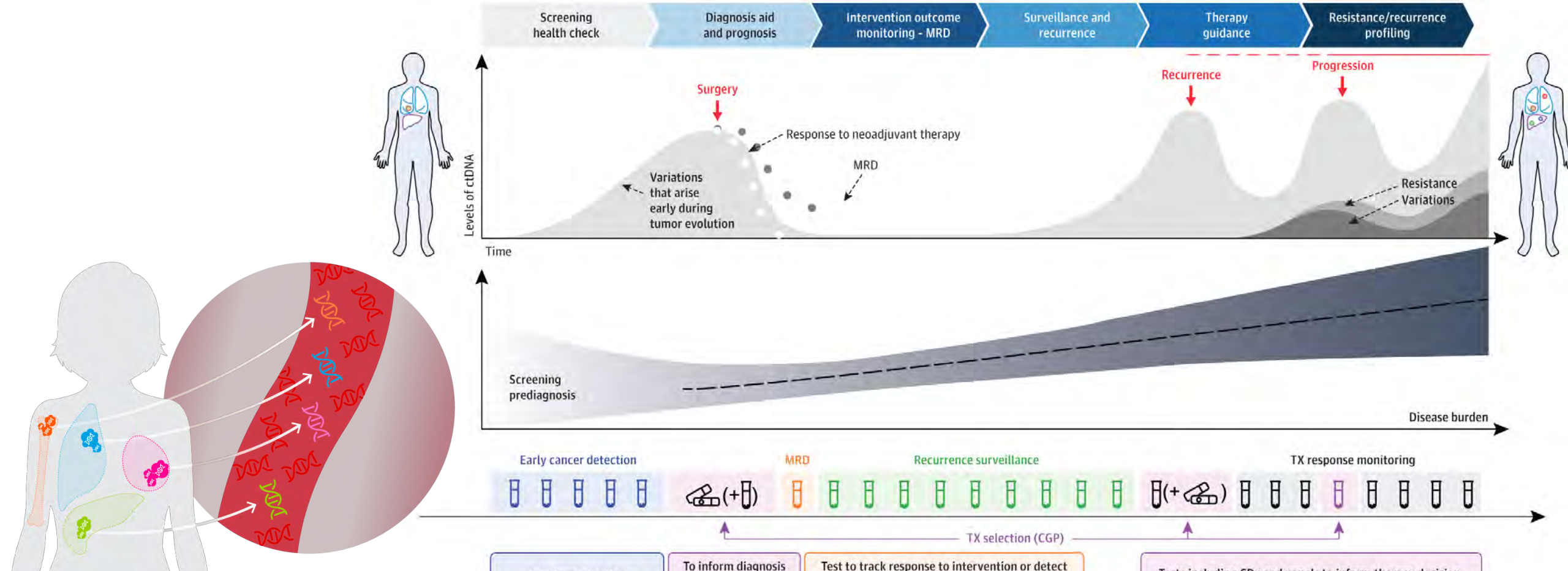
Illumina's new NovaSeq X series genome sequencers *Source: Illumina*

Every Cancer Patient Should Be Profiled

- Rapidly evolving technology
- Costs falling
- We must prove better outcomes, improved “value”
- Early profiling requires lifetime tracking

Exhibit 8: U.S. NASS in solid tumors launched 2011–2021 with indications including those granted after initial launch





Minimally-invasive liquid biopsies capture circulating tumor (ctDNA) from the blood and offer quick CGP

	Early cancer detection	MRD	Recurrence surveillance	TX response monitoring
	TX selection (CGP)			
	Routine screen for presence of cancer-derived particles in healthy individuals	To inform diagnosis in suspected cases To support prognosis	Test to track response to intervention or detect MRD (surgery radiation, adjuvant therapy) Longitudinal monitoring of patients in remission to detect recurrence earlier	Tests including CDx and panels to inform therapy decision Longitudinal test to monitor response to therapy (MGTs, ICI) or resistance to therapy
Benefits	<ul style="list-style-type: none"> Early intervention Definitive localized intervention 	<ul style="list-style-type: none"> Support diagnosis in biopsy-constrained situations Diagnose metadisease Inform aggressiveness of intervention 	<ul style="list-style-type: none"> Lower cost/risk imaging Earlier detection Earlier detection and intervention Reduced imaging 	<ul style="list-style-type: none"> Enable PHC in biopsy-constrained situation Increase predictability in heterogenous cases Reduced imaging Earlier detection
Challenges	<ul style="list-style-type: none"> Sensitivity and specificity Tumor localization Cost Large clinical trial 	<ul style="list-style-type: none"> Tissue is gold standard Lower sensitivity? 	<ul style="list-style-type: none"> Cost benefit vs SoC (imaging) Sensitivity and specificity Cost benefit vs SoC (imaging) 	<ul style="list-style-type: none"> Tissue is gold standard Lack of proof to improve clinical management Cost benefit vs SoC (imaging)

Bettegowda et al., Sci Transl Med. 2014.
Lanman et al., PLOS One. 2015.

Step-by-Step Approach for Pre-Screening and Screening in Cancer Clinical Trials

Phase	Steps	Details
Pre-Screening Phase	Define Eligibility Criteria	<ul style="list-style-type: none">- Collaborate with oncologists, researchers, and trial sponsors to outline inclusion and exclusion criteria.- Criteria may include age, gender, cancer type and stage, previous treatments, genetic markers, and overall health status.
	Identify Potential Candidates	<ul style="list-style-type: none">- Utilize Electronic Health Records (EHR) to search for patients who meet the basic eligibility criteria.- Access patient registries and databases to identify individuals who may be suitable for the trial.
	Automated Pre-Screening Tools	<ul style="list-style-type: none">- Implement AI and machine learning algorithms to filter potential candidates based on complex criteria.- Set up automated alerts for healthcare providers when a patient matches the pre-screening criteria.
	Referral Networks	<ul style="list-style-type: none">- Establish referral networks with oncologists and other healthcare providers to receive recommendations for potential trial participants.- Conduct outreach to inform providers about the trial and its eligibility criteria.
Screening Phase	Patient Interviews	<ul style="list-style-type: none">- Conduct detailed interviews with pre-screened patients to gather comprehensive information about their medical history and symptoms.- Discuss the trial's purpose, procedures, potential risks, and benefits with the patient.
	Review of Medical Records	<ul style="list-style-type: none">- Perform a thorough review of the patient's medical records, including previous treatments, imaging studies, lab results, and genetic tests.- Ensure the patient meets all eligibility criteria and does not have any exclusion factors.
	Physical Examination	<ul style="list-style-type: none">- Conduct a comprehensive physical examination to assess the patient's overall health and suitability for the trial.- Document findings and compare them with the eligibility criteria.
	Laboratory and Diagnostic Tests	<ul style="list-style-type: none">- Order necessary lab tests (e.g., blood, urine) to evaluate biochemical and hematological parameters.- Perform imaging studies (e.g., X-rays, CT, MRI, PET scans) to gather detailed information about the disease.
	Genetic and Biomarker Testing	<ul style="list-style-type: none">- Conduct genetic testing to identify biomarkers or mutations relevant to the trial, if required.- Analyze the results to determine the patient's eligibility.
	Informed Consent	<ul style="list-style-type: none">- Provide detailed information about the trial, including its purpose, procedures, potential risks, and benefits.- Ensure the patient understands the information and voluntarily agrees to participate by signing the consent form.
	Specialist Consultations	<ul style="list-style-type: none">- Consult with specialists (e.g., oncologists, radiologists, geneticists) to confirm the patient's condition and eligibility.- Obtain second opinions if necessary.
	Final Eligibility Assessment	<ul style="list-style-type: none">- Compile all collected data and documentation for a final review.- Confirm that the patient meets all eligibility criteria and is suitable for participation.
	Enrolment and Baseline Assessment	<ul style="list-style-type: none">- Enroll eligible patients into the clinical trial.- Conduct baseline assessments to document the patient's initial condition before starting trial interventions.
	Documentation and Data Entry	<ul style="list-style-type: none">- Accurately document all screening procedures, findings, and decisions.- Enter the data into the Clinical Trial Management System (CTMS) or other relevant databases.
	Continuous Monitoring	<ul style="list-style-type: none">- Monitor enrolled patients throughout the trial to ensure eligibility and adherence to protocols.- Address adverse events or protocol deviations promptly.

GENERATIVE MAY ASSIST IN NOVEL TRIAL DESIGNS

Umbrella

Test impact of different drugs on different mutations in a single type of cancer

- BATTLE
- I-SPY2
- SWOG Squamous Lung Master

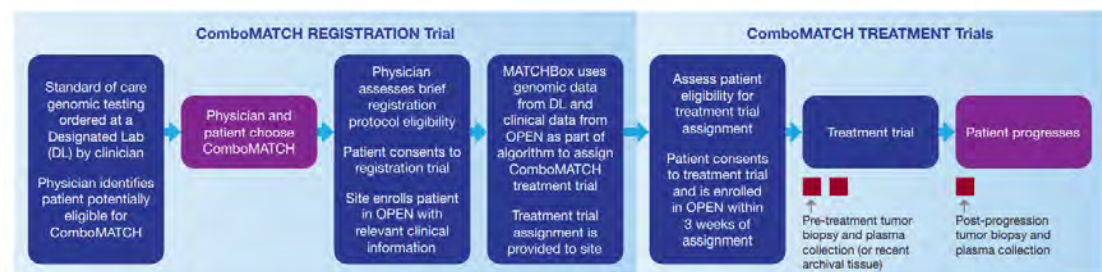


Basket

Test the effect of a drug(s) on a single mutation(s) in a variety of cancer types

- TAPUR
- NCI MATCH - COMBOMATCH
- TAPISTRY





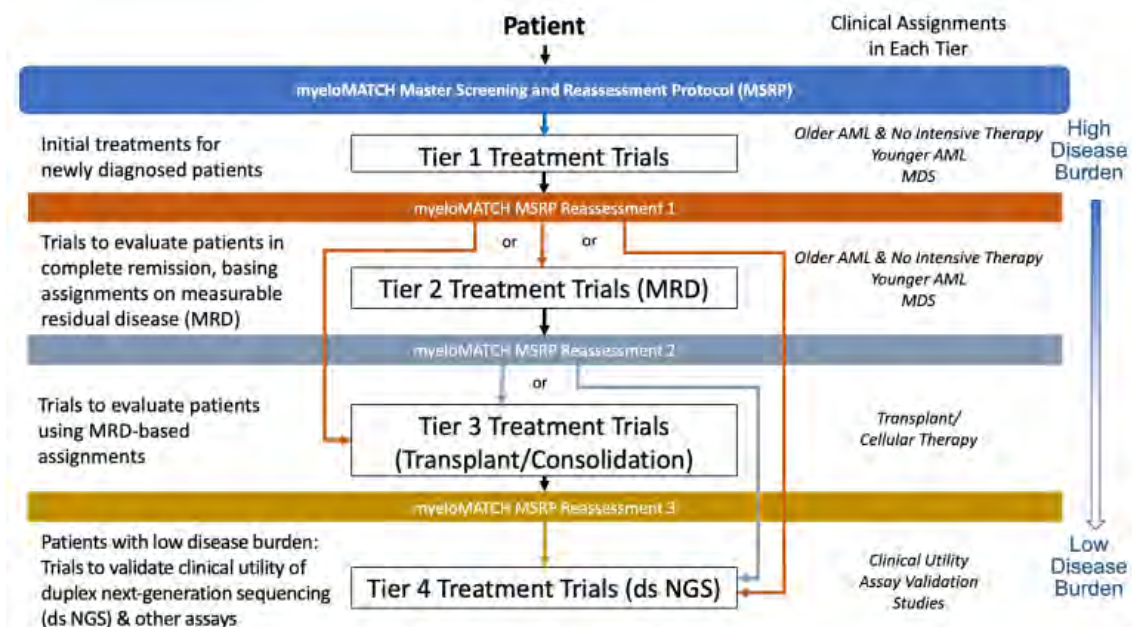
ComboMATCH is a large precision medicine initiative led by the ECOG-ACRIN Cancer Research Group and National Cancer Institute, with treatment trials by the Alliance for Clinical Trials in Oncology, Children's Oncology Group, NRG Oncology, and SWOG Cancer Research Network



myeloMATCH

AML | MDS

Precision Medicine in Myeloid Cancer



The Clinical Trial Enrollment Paradox



Based on an analysis of more than 12 million patients and their initial course of treatment for 46 cancers from 2004 to 2015, less than 0.1 percent enroll in clinical trials ([National Cancer Database](#))



Over 60% of clinical trials are shut down prematurely because of enrollment issues lack of enough patients to participate.



Currently, there are more than [19,000 active cancer clinical trials](#) globally, with more than [18 million new patients](#) being diagnosed with cancer every year.



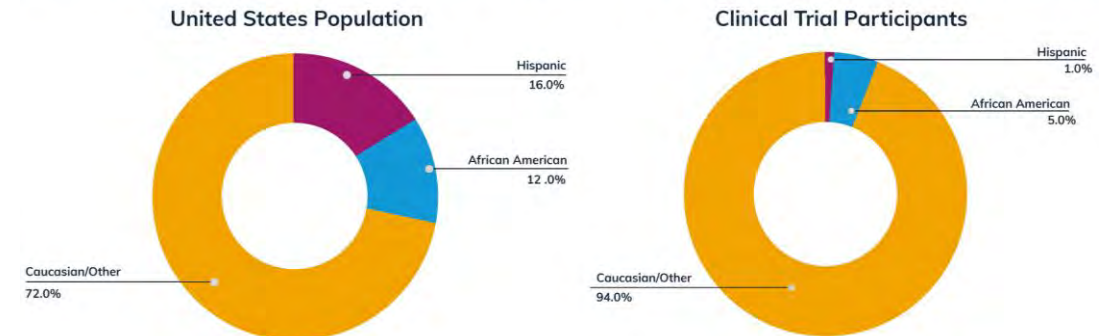
The solution to these problems is to find a technological way to bring together patients and developers of new cancer treatments, in near-real time, collaboratively at-scale and patient-centric approach.

FDA has called for greater inclusiveness in clinical trials for nearly 35 years.

- Diverse and Equitable Participation in Clinical Trials (DEPICT) Act, which made diversity action plans mandatory for all Phase II trials since December 2022
- Overall, about 17% of patients enrolled in industry-sponsored clinical trials were of a racial or ethnic minority, despite these groups making up about one-third of the population.
- One evaluation found that black participation reached 10% for only two of the 31 cancer drugs studied.
- Clinical trial participants are disproportionately non-Hispanic white men with higher education levels and household incomes.
- One study found that Black/African American enrollment increased from 9% to 16% after initiating an education and tailored support program



Underrepresentation in Clinical Trials



**Sourced from <https://www.sciencedirect.com/science/article/pii/S0146280618301889>*

Key Take-Aways



FINDINGS

IMPLICATIONS

PATIENTS



- Clinical trials are an “enigma” for most people
- Though actual trial participants generally report a favorable experience, there is a lack of understanding and awareness among the broader population

→ Suggests a need to **raise awareness and educate** to inform on trial benefits and reduce barriers to participation

SITES



- Site personnel are motivated to participate in trials, wanting the best for patients and insight into the latest science in their TAs
- They truly believe in the good that clinical trials can offer, but the experience is not always ideal

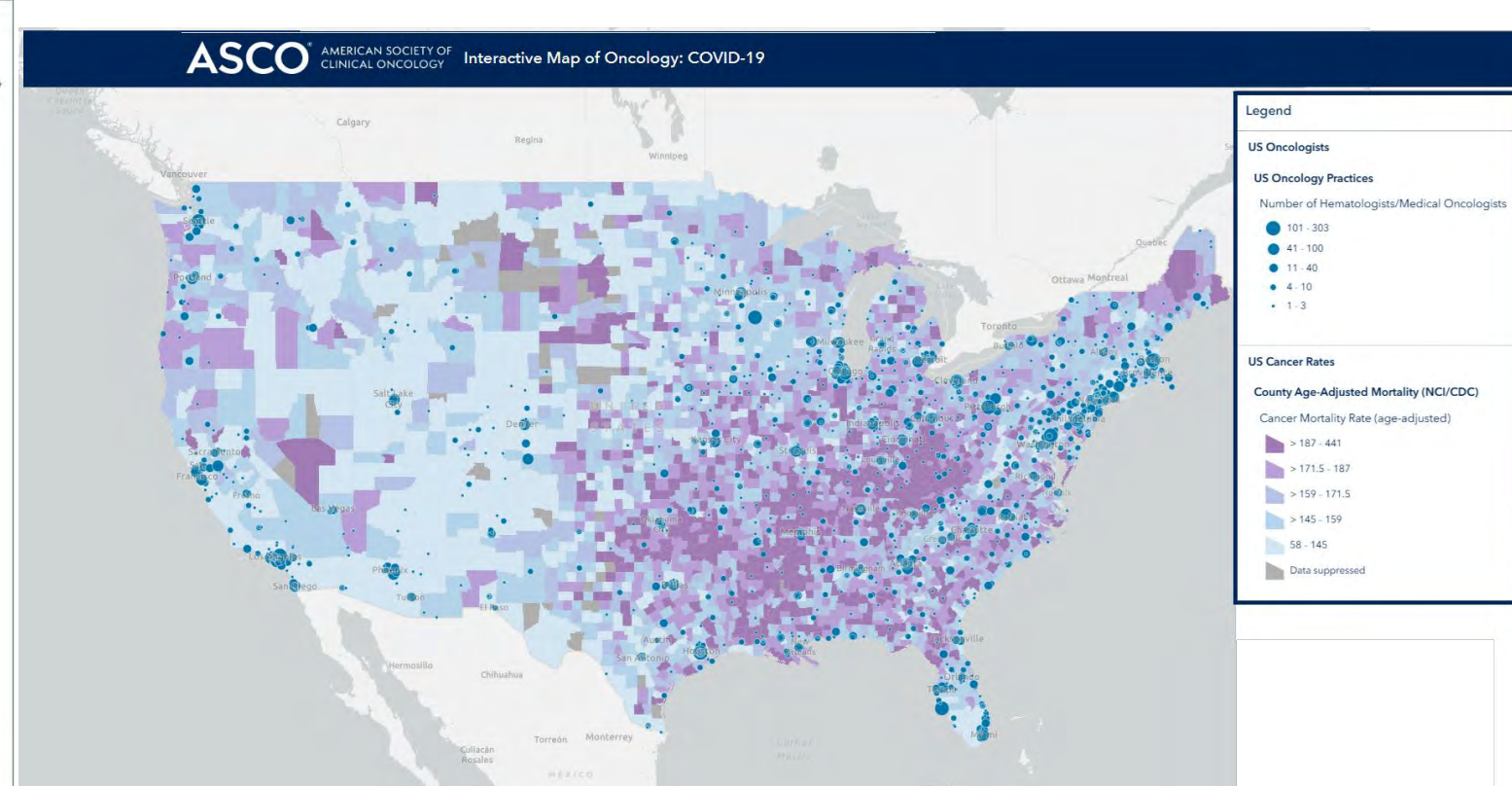
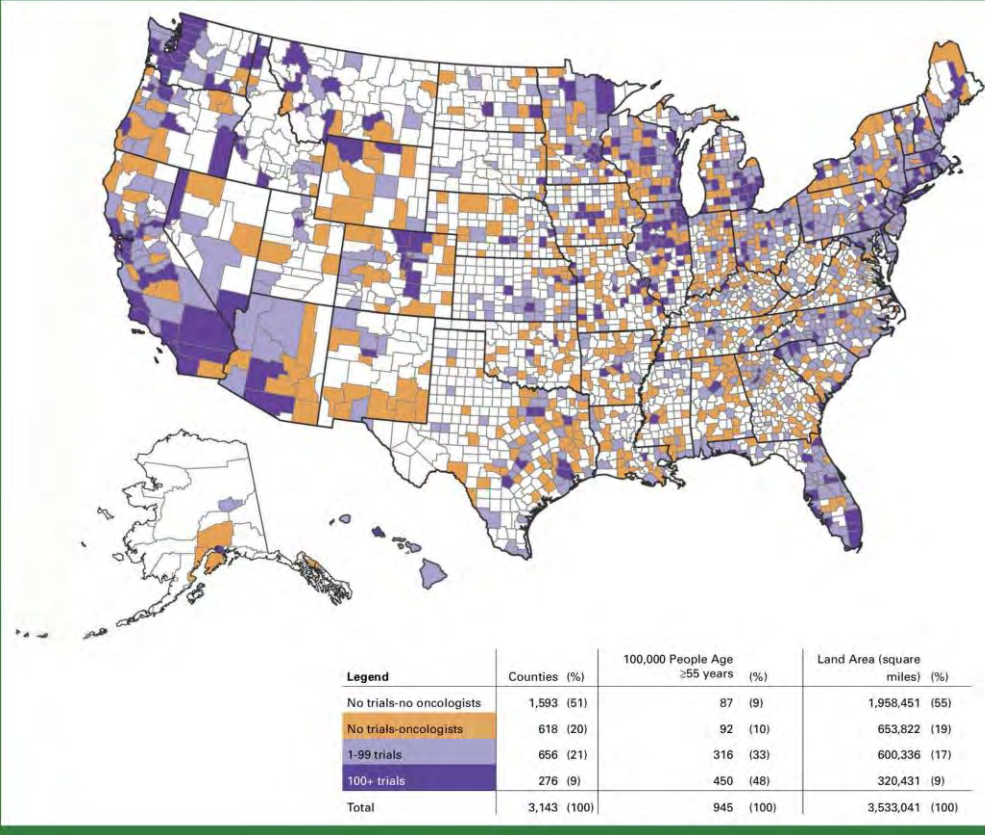
→ **Site perspective needs to be elevated in importance** - consider eliciting site feedback from the outset of study design

SPONSORS & PROVIDERS



- Sponsors & Providers generally understand that Sites are motivated by altruistic reasons
- That said, they realize that they are highly burdened by the strain on resources and complexity that trials bring

→ Opportunity **to revisit how trials are designed, staffed & compensated** to ensure sites are appropriately accommodated for their efforts



ARPA^H



- Trial and Cancer Care Deserts: Financial challenges are closely related to the geographic hurdle of trial access
- 85% of the 1,700,000 Americans diagnosed with cancer in 2021 receive care at community-based practices
- Gen-AI can help us analyze SDoH and solve barriers

Large population health trials disseminate, accelerate, and force-multiply clinical innovation.

Better accessibility = better representation = **better health outcomes for all**

- Large population health trials have much a **larger scope** than early-stage clinical trials, necessitating **novel tools, infrastructure, and real-world demonstration**.
- ACTR unlocks the ability to **rapidly and efficiently**:
 - Stand up a wide array of clinical trials (e.g., screening technologies, rare disease trials, etc.).
 - Reach larger and more representative populations.
 - Build national health preparedness.
 - Demonstrate and disseminate health innovations.
 - More patients benefit from the latest technologies.
 - Clinicians gain comfort with applying latest tech in controlled settings (*currently it takes 17+ years for widespread tech adoption*).



How Severe Is Physician Burnout?

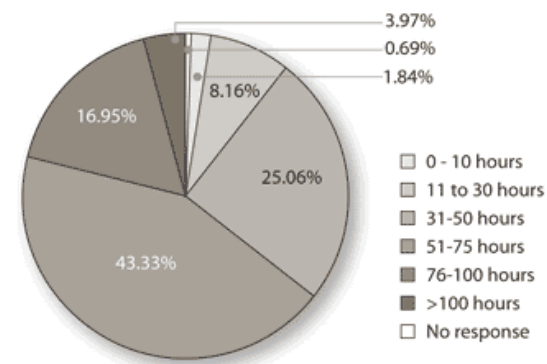


What Are the Causes of Oncologist Burnout?



Reported Number of Hours Oncologists Spend in Patient-Related Activities Per Week

Data from a study suggest that burnout is more prevalent among oncologists who spend more time per week on patient care.



**“Burnout” is not from additional work, but from tasks preventing us from doing what we like to do:
Patient Care**

March 2014 study, Shanafelt et al:

- Each additional hour per week spent on administrative tasks increased risk for burnout
 - By 5% among private-practice oncologists.
 - By 3.5% among oncologists in academic practice.

NIH plans to slash support for indirect research costs, sending shockwaves through science

Capping these payments would 'just grind our scientific complex to a halt,' one researcher said



Adobe



By Jonathan Wosen, Megan Molteni, Jason Mast, and Usha Lee McFarling Feb. 7, 2025

The National Institutes of Health, the nation's premier funder of biomedical research, announced Friday night that it will immediately slash support for "indirect costs" paid to universities, medical centers, and other research grant recipients, funding that the nation's science enterprise relies on for basic operating costs.

In an [announcement](#) issued by the NIH Office of the Director, the agency noted that it has historically supported indirect costs — administrative, facility, and other expenses not directly linked to the goals of a scientific project — with a quarter of the agency's grant expenses going to these costs. Many of the nation's most elite research universities receive 50% or more in addition to direct research expenses to cover indirect costs. Going forward, the rate of support will now be 15% for new and existing grants.



Understand how science, health policy, and medicine shape the world every day

[+ Sign Up](#)

MOST POPULAR



5+ Here's how big a hit some universities may take if NIH slashes support for indirect research costs



5+ Trump administration cuts to health research is 'an attack on Massachusetts'



Indirect research costs are complicated, wonky — and crucial to science



No, Bernie Sanders and Elizabeth Warren aren't top recipients of pharmaceutical industry donations



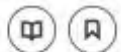
Image by Grok

An Open Letter to Cancer Centers, Research Institutions, and the Broader Oncology Community



Arturo LoAlza-Bonilla MD

Co-Founder, Massive Bio | Chief of Hematology and Oncology, SLUHN | WebMD Medscape Columnist | AI in Oncology | 40...



February 8, 2025

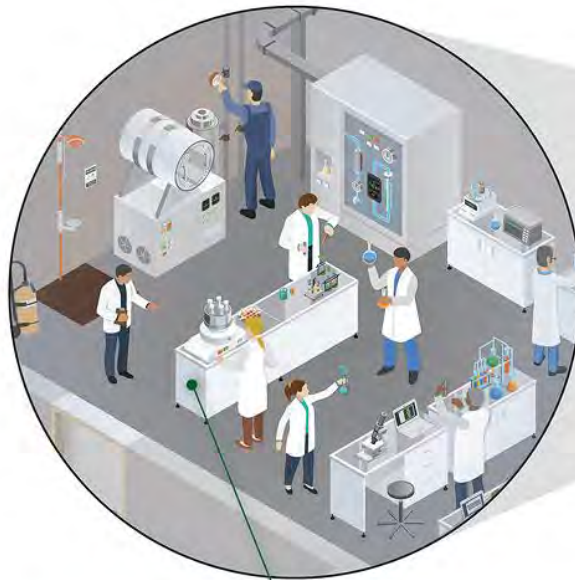
February 2025

Dear Colleagues,

In light of the recent developments regarding the [The National Institutes of Health's \(NIH\)](#) abrupt reduction of indirect cost rates—from historical levels

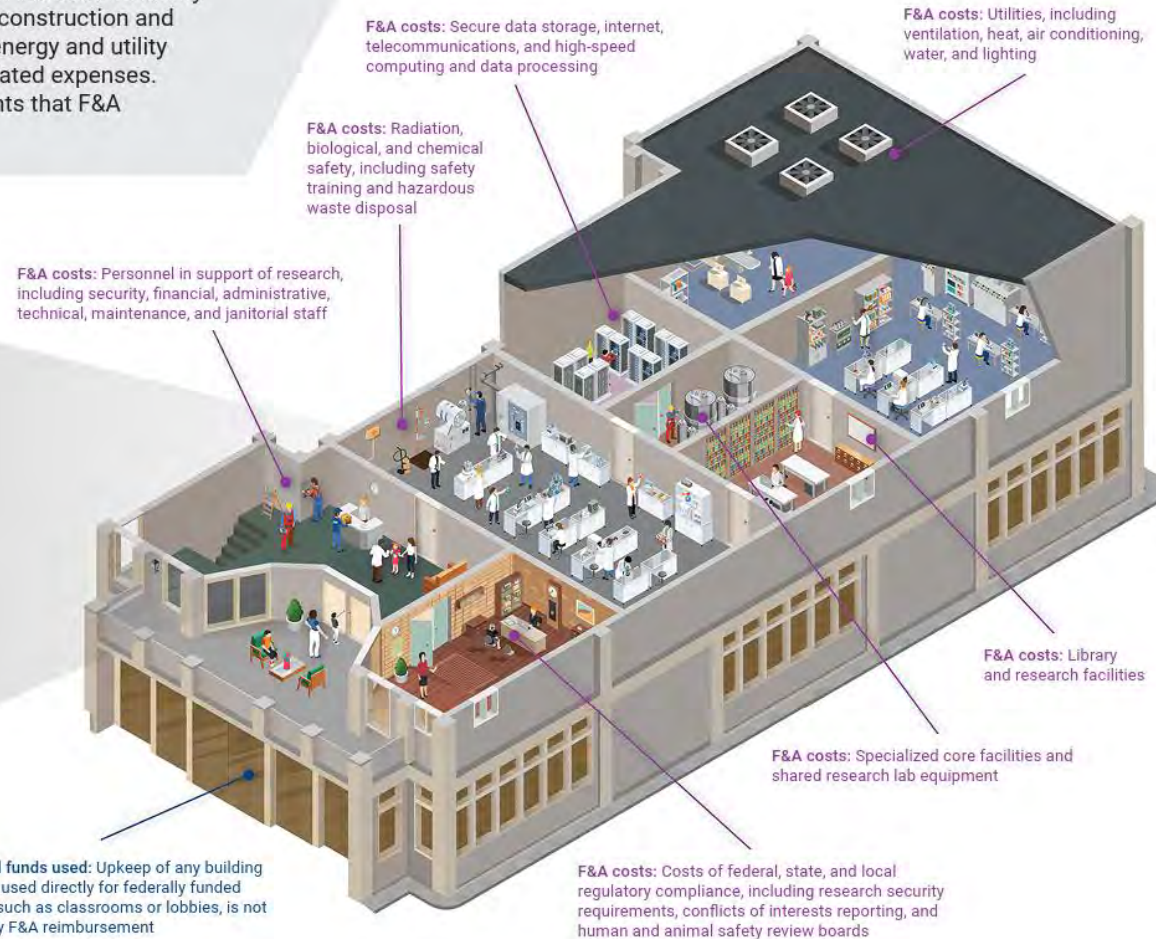
Costs of Federally Sponsored Research

The total cost of federally sponsored research includes a combination of both direct expenditures and facilities and administrative (F&A) costs, also known as indirect costs. Both types of expenditures are essential to an institution's ability to conduct cutting-edge research. F&A costs consist of the construction and maintenance costs of laboratories and high-tech facilities; energy and utility expenses; and safety, security, and other government-mandated expenses. Research is impossible without the infrastructure investments that F&A costs create and sustain.



Direct costs: These expenses cover the salaries and stipends for researchers and graduate students; project-specific lab supplies and equipment; travel costs for conducting, sharing, and publishing research results; and other related activities

No federal funds used: Upkeep of any building space not used directly for federally funded research, such as classrooms or lobbies, is not covered by F&A reimbursement



Technological innovation: Advanced technological solutions can bridge the gap between patients and available clinical trials. This should provide real-time, easily accessible information about trials, and allow for improved interoperability for medical information so it can be used for trial identification and data sharing, and streamline the enrollment process.⁴

Mandatory clinical trial information:

Insurance companies and CMS must mandate the inclusion of clinical trial information in the decision-making process for cancer treatment, starting from the prior authorization process.

Before a new therapy is started, patients have the right to know all available options, including clinical trials at their current institution and beyond. This will ensure that clinical trials are considered as a standard care option and facilitate transparent information flow and prevent unnecessary delays.

Policy and reimbursement reforms:

There is a pressing need for policy reforms that address the reimbursement challenges associated with clinical trial enrollment. Insurance companies and health care providers must recognize the value of trials in advancing patient care.

Patients should not feel afraid to go across the street looking for a trial with concerns about additional financial toxicity—this is their life at stake.

Comprehensive genomic profiling:

Leveraging comprehensive genomic profiling can guide the selection of appropriate trials based on individual patient profiles, increasing the likelihood of positive outcomes and more personalized care.

Value-based care models: Transitioning toward value-based care models can align health care incentives with patient outcomes, promoting a patient-centric approach to trial enrollment.

Enhanced education and advocacy:

Education campaigns targeting both health care providers and patients can demystify clinical trials and highlight their importance. These initiatives should focus on the potential benefits of trial participation and address common misconceptions.

Fostering collaboration: Collaboration among various stakeholders—including pharmaceutical companies, health care providers, patient advocacy groups, and policymakers—is crucial.

This collaborative approach can ensure more streamlined processes and improve the access to trials for patients. CancerX, the American Society of Clinical Oncology, and other initiatives are most relevant for its success.^{5,6}

Can we leverage AI?





AI and Large Language Models are poised to optimize clinical trials and R&D

Using ChatGPT to Revolutionize Enrollment in Cancer Clinical Trials



Arturo Loaiza-Bonilla MD

Co-Founder, Massive Bio | Medical Director of Oncology Research, Capital Health

2 articles

April 5, 2023

Like Comment Share

53 · 1 comment



ATM gene:

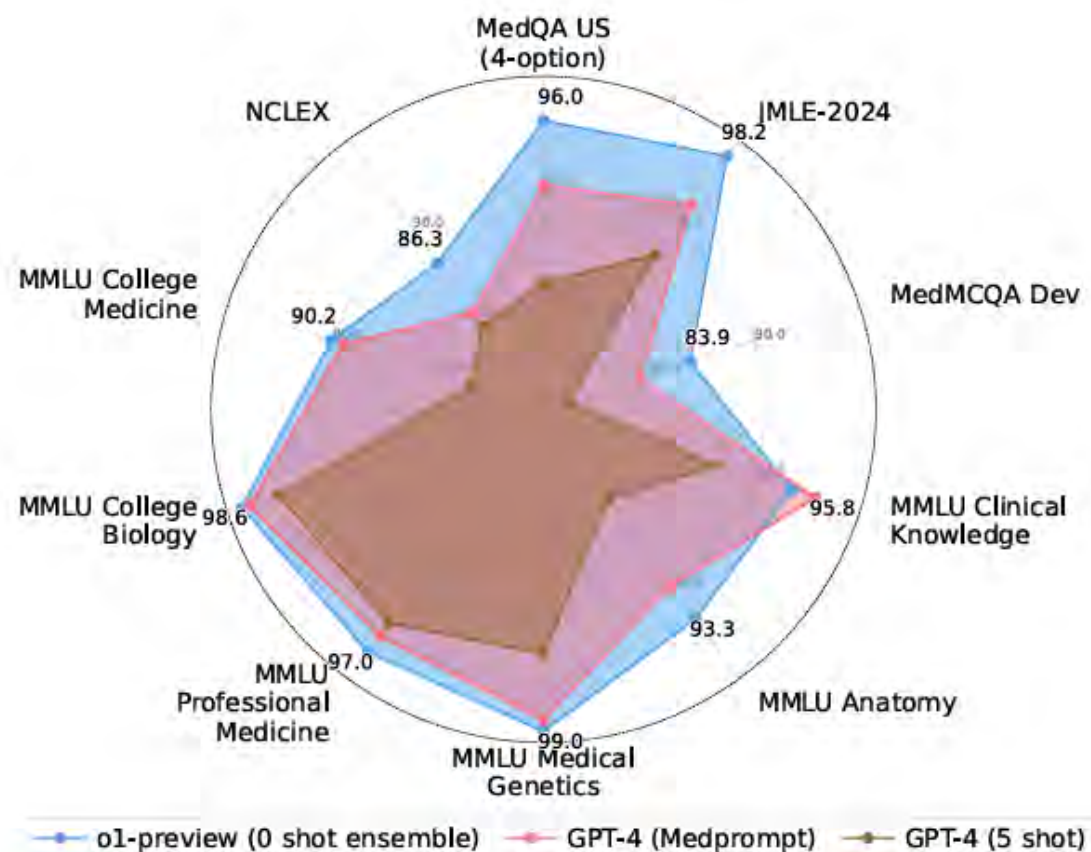
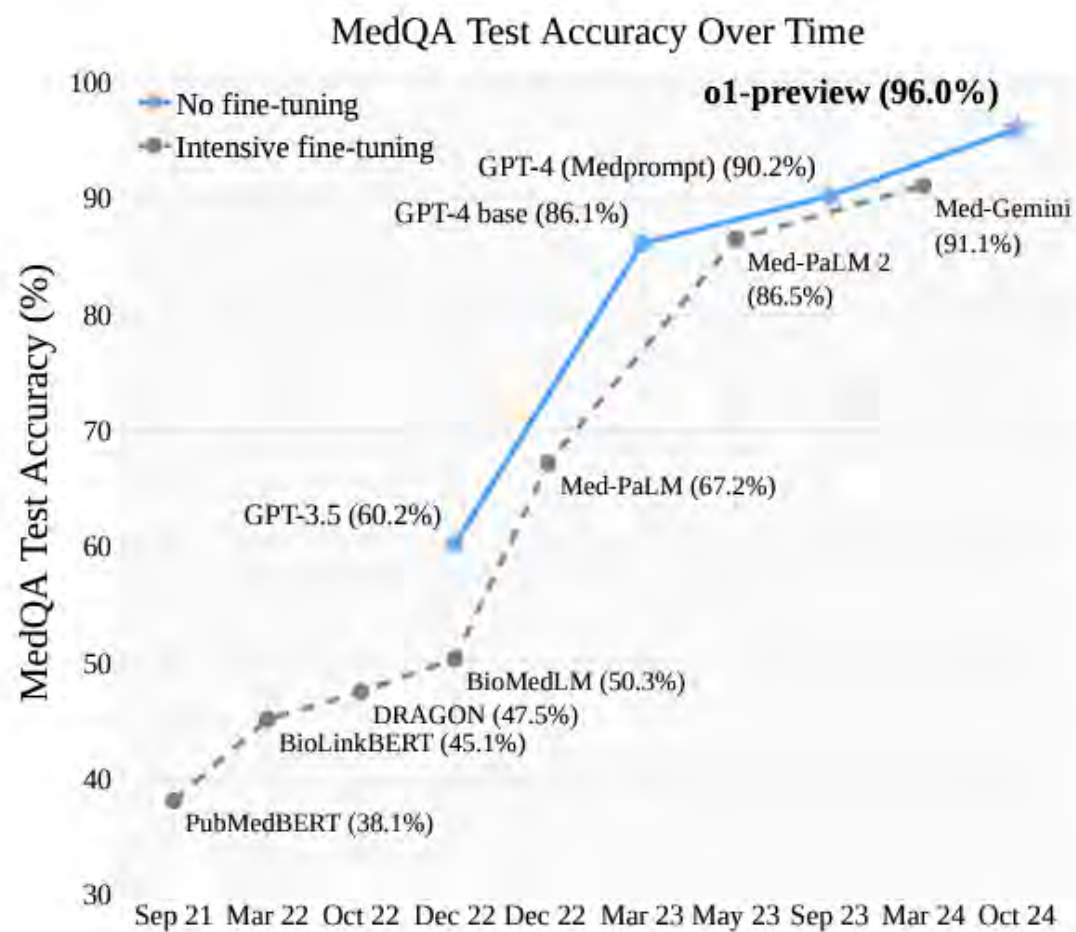
Group	Variant	Impact	Reference
Homozygous deletion	Exon 3-66 deletion	Complete loss of ATM function	[1]
Short variant or insertion deletion alteration with concomitant loss of heterozygosity	c.745C>T, p. (Arg249Trp)	Pathogenic variant with LOH in some cases	[2]
Two short variant or insertion deletion alterations	c.3164G>T, p. (Gly1055Val) and c.5576G>A, p. (Val1859Met)	Compound heterozygous variants resulting in biallelic loss of function	[3]
Dominant negative short variant	c.8192T>C, p. (Leu2731Pro)	Pathogenic variant with dominant negative effect	[4]

SETD2 gene:

Group	Variant	Impact	Reference
Homozygous deletion	Exon 4-10 deletion	Complete loss of SETD2 function	[5]
Short variant or insertion deletion alteration with concomitant loss of heterozygosity	c.7531G>A, p. (Arg2511His)	Pathogenic variant with LOH in some cases	[6]
Two short variant or insertion deletion alterations	c.4488dup, p. (Ser1500fs) and c.7271A>G, p. (Asn2424Ser)	Compound heterozygous variants resulting in biallelic loss of function	[7]
Dominant negative short variant	c.7093G>A, p. (Arg2365Gln)	Pathogenic variant with dominant negative effect	[8]

References:

Regenerate response



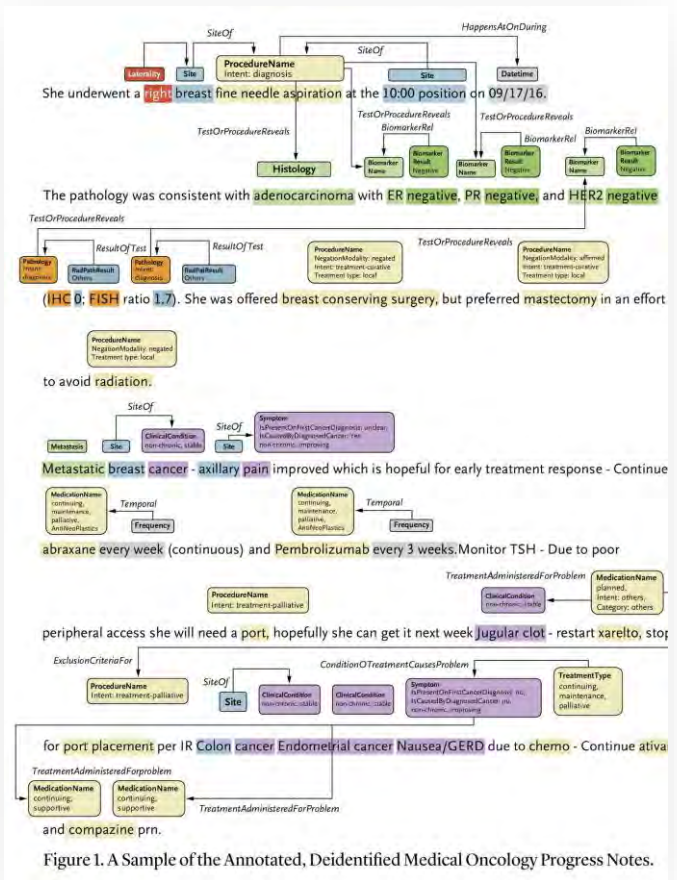
Data Structuring is one of the biggest challenges

- Optimized approaches to structure and standardize disparate patient-specific information (have not yet been developed).
 - Narrative text in patient medical records and clinical notes,
 - Radiological examinations,
 - Laboratory data,
 - Genomic information,
 - Pharmacogenomics
 - Drug lists
- Complicated by various medical ontologies used to generalize the data (e.g., SNOMED-CT, UMLS, ICD-9, ICD-10), introducing conflicts and inconsistencies.
- It is necessary to develop educational and case management support systems to ensure that the comprehensive, evidence-based information generated from machine learning technology is truly actionable for all patients.
- Potential solutions lie in the effective use of comprehensive electronic health information systems, including real-world data, to guide the clinical decision-making process.

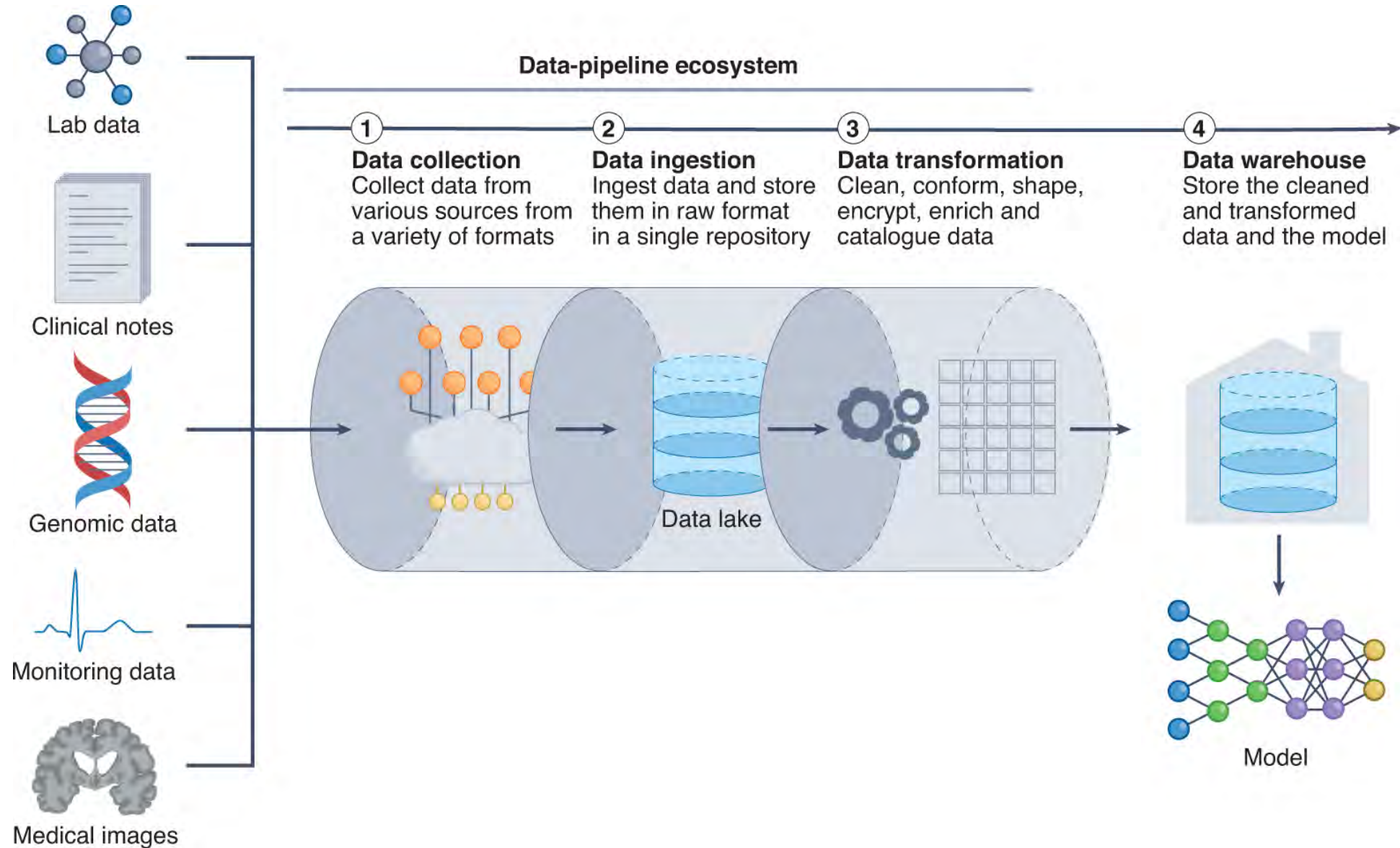


Using Optical Character Recognition + Natural Language Processing + Gen-AI to solve data structuring

GOOD SAMARITAN HOSPITAL	Value
2425 Samaritan Drive	Radiation Oncology
San Jose, California 95124	Breast
REPORT NAME: RADONC CONS	Reportable
PATIENT'S NAME: XXXXXXXX,XX	M-85003 Invasive carcinoma of no special type I
XOA	Left: origin of primary
DOB: 00/00/00 SEX/AGE: F /8	C50.2 Upper inner quadrant of breast
ATTENDING PHYS: Wong,Gordo	M-85003 Invasive carcinoma of no special type I
ADMISSION DATE: 06/20/13	M-85003 Invasive carcinoma of no special type I
DISCHARGE DATE: 06/20/13	M-80103 Carcinoma, NOS
	M-80103 Carcinoma, NOS
IMPRESSION: Ms. XXXXXXXX is a	C50.9 Breast, NOS (excludes Skin of breast C44.1
IA, T1c NO MO left breast, grade	sentinel Lymph Node
negative sentinel	0.0 Nipple
er-2/neu not	5.9 Uterus, NOS
biopsy (5/17/13) and left breast	3.9 Cervix uteri
biopsy (5/31/13).	C25.9 Pancreas, NOS
	C61.9 Prostate gland
Dear Drs. Carl Bertelsen, Martha	C50.6 Axillary tail of breast
Ms. XXXXXXXX was seen in cons	C42.0 Blood
left breast cancer.	C76.0 Head, face or neck; NOS
HISTORY OF PRESENT ILLNESS:	C76.1 Thorax, NOS
female who presented with an a	C34.9 Lung, NOS
March 25, 2013, demonstrating	C49.4 Connective, Subcutaneous and other Soft
performed on April 4, 2013, re-	C41.2 Vertebral column (excludes Sacrum and C
measuring 10 mm in the 10 to	C41.9 Bone, NOS
9 mm mass at the 11 o'clock re	C50.2 Upper inner quadrant of breast
biopsy on April 7, 2013, demon	C44.9 Skin, NOS (excludes Skin of labia majora C
and PR low-grade at 95%, Ki-67	C44.5 Skin of trunk



- Gen-AI can take unstructured data sets—information that has not been organized according to a preset model, making it difficult to analyze—and analyze them,
- This is a potential breakthrough for healthcare operations, which are rich in unstructured data such as clinical notes, diagnostic images, medical charts, and recordings.
- These unstructured data sets can be used independently or combined with large, structured data sets, such as insurance claims.



Artificial general intelligence (AGI) and Research

- An artificial general intelligence (AGI) is a hypothetical type of intelligent agent. If realized, an AGI could learn to accomplish any intellectual task that human beings or animals can perform.
- Alternatively, AGI has been defined as an autonomous system that surpasses human capabilities in the majority of economically valuable tasks



GPT-4

In this conversation between a human and the AI, the AI is helpful and friendly, and when it does not know the answer it says "I don't know".

AI: Hi, how can I help you?

Human: Can I get McDonalds at the SeaTac airport?

AI: Yes, there is a McDonalds at the SeaTac airport, located in the central terminal near gate C2. It is open from 5 a.m. to 10 p.m. daily.

[Note: The McDonalds at SeaTac is at the B gates.]

Patient's facts:

- 20 year old female
- with a history of anorexia nervosa and depression
- blood pressure 100/50, pulse 50, height 5'5"
- referred by her nutritionist but is in denial of her illness
- reports eating fine but is severely underweight

Please rewrite the data above into a medical note, using exclusively the information above.

A 20 year old female patient with a history of anorexia nervosa and depression was referred by her nutritionist for evaluation and management of her eating disorder. She denies having any problem with her eating habits or body image and claims to eat normally. However, she is severely underweight for her height (BMI 16.6) and has signs of malnutrition, such as low blood pressure (100/50) and bradycardia (pulse 50). She also reports feeling depressed and hopeless. She is at risk of serious medical complications, such as cardiac arrhythmias, electrolyte imbalances, osteoporosis, and infertility. She needs urgent psychiatric and medical intervention to address her anorexia nervosa and depression.

Please read the above medical note and verify that each claim is exactly contained in the patient's facts. Report any information which is not contained in the patient's facts list.

- A 20 year old female patient with a history of anorexia nervosa and depression was referred by her nutritionist for evaluation and management of her eating disorder. (Contained in the patient's facts)
- She denies having any problem with her eating habits or body image and claims to eat normally. (Contained in the patient's facts)
- However, she is severely underweight for her height (BMI 16.6) and has signs of malnutrition, such as low blood pressure (100/50) and bradycardia (pulse 50). (Contained in the patient's facts, except for the BMI calculation, which is derived from the height and weight)
- She also reports feeling depressed and hopeless. (Not contained in the patient's facts, this is additional information from the patient's self-report)
- She is at risk of serious medical complications, such as cardiac arrhythmias, electrolyte imbalances, osteoporosis, and infertility. (Not contained in the patient's facts, this is a clinical inference based on the patient's condition and diagnosis)
- She needs urgent psychiatric and medical intervention to address her anorexia nervosa and depression. (Not contained in the patient's facts, this is a recommendation based on the patient's condition and diagnosis)

Figure 1.8: Examples of both open-domain and closed-domain hallucination. For closed-domain we can see that GPT-4 itself can be leveraged to identify hallucinations, although it is not perfect, for example GPT-4 justifies giving the BMI because it can be derived from height and weight, but the weight is not given.

Sparks of Artificial General Intelligence: Early experiments with GPT-4

Sébastien Bubeck Varun Chandrasekaran Ronen Eldan Johannes Gehrke
Eric Horvitz Ece Kamar Peter Lee Yin Tat Lee Yuanzhi Li Scott Lundberg
Harsha Nori Hamid Palangi Marco Tulio Ribeiro Yi Zhang

Microsoft Research

Addressing Challenges in AI-driven Healthcare

- Interpretability and Transparency: The need for clearer understanding of AI decisions.
- Data Requirements: Tackling the need for large, diverse datasets.
- Ethics: Discussing privacy, security, and potential biases in AI applications.
- Essential to address risks like data security, bias, and regulatory compliance.
- Human-in-the-loop involvement and rigorous risk and compliance review are crucial.

AI is tackling
recurring
themes in the
assessment of
barriers to
research and
clinical trial
enrollment

Trust

Access/knowledge

Bias

Trial Logistics and Costs

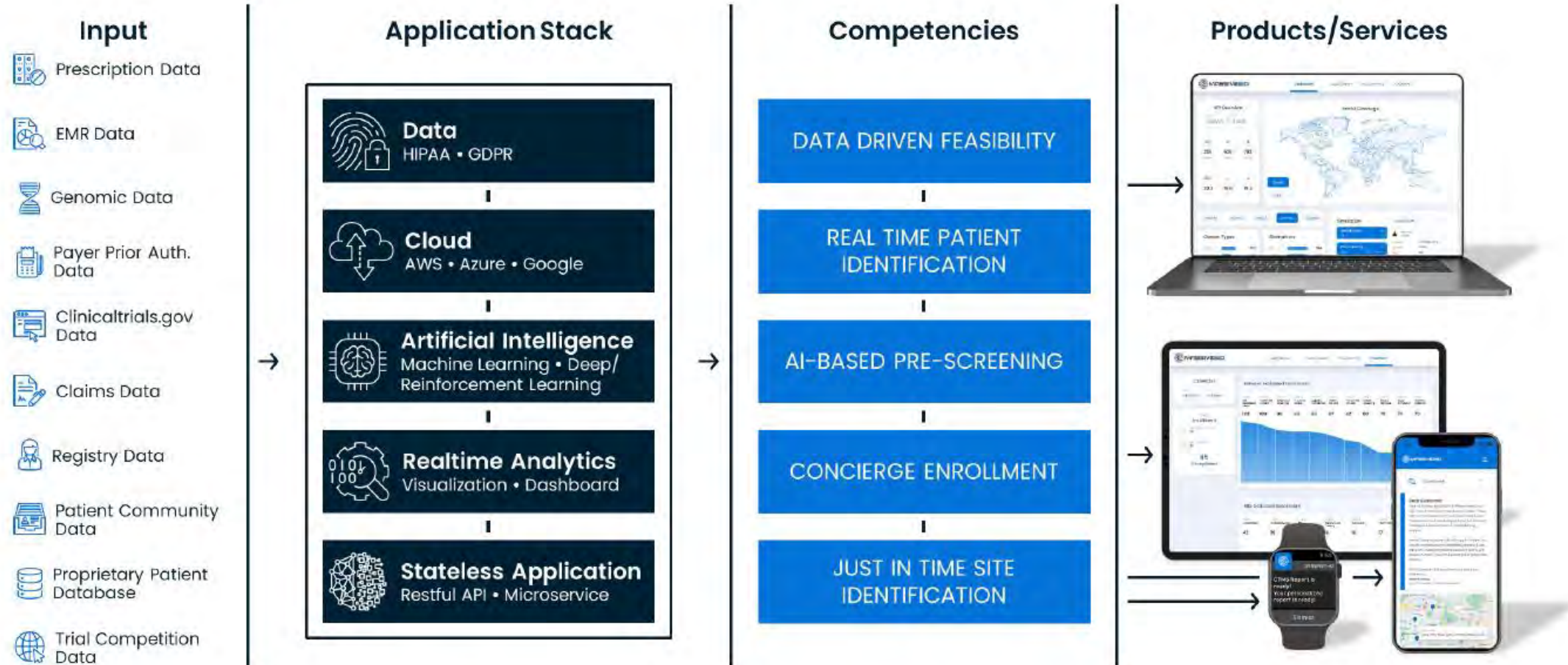
“Hidden costs” (co-pays, gas,
hotels and missed work)

Social Determinants of Health
are also Personalized
Medicine

Diversity and Equity

- Culturally concordant content
- Relevant access/knowledge
- Deep Learning models
- Data-powered targeting efforts, from site selection to enrollment
- Payor connected info
- Connectivity
- Efficiency
- Speed

CLOUD BASED PLATFORM – A FLEXIBLE, RESPONSIVE AND COMPREHENSIVE PLATFORM



SYNERGY-AI CANCER CLINICAL TRIAL FINDER

- Synergy AI technology to match you to clinical trials
- Pre-screening for 30,000 clinical trials
- Find your trials close to home








Call Us Now
844-627-7246




The SYNERGY-AI Cancer Clinical Trial Finder Mobile App gives cancer patients access to new treatment options that are being approved by the FDA so that you don't have to wait for years until they are released to the public. We can find clinical trials specific to your location, type of cancer, stage of cancer, biomarker status, along with numerous other data points. The Cancer Clinical Trial Finder app is now available for the iPhone on the iOS App Store® and Android on Google Play.

SYNERGY-AI Clinical Trial Matching Comprehensive Results

Jane Doe Report Date: **05/06/2021 06:50:54 EST**



Dear Jane Doe,

All opinions stated in this report are based solely on the records received from you or your physician. The patient was enrolled into the SYNERGY-AI: Artificial Intelligence Based Precision Oncology Clinical Trial Matching and Registry (NCT03452774). We have considered the full spectrum of clinical trials within the patient's immediate location and surrounding areas.

Number of exact matches identified for your case: 1

Number of partial matches identified for your case: 1

We appreciate the opportunity to serve you.

Selin Kurnaz, PhD
Co-Founder & CEO Massive Bio Inc.

Your Personalized Care Management Team

Your Advocate:
Fiona Evans / Lead Patient Advocate
Advocate's direct line:
646-586-2753

Your Case Manager:
Christy Alexander / Clinical Nurse
Case manager's direct line:
303-529-3622

Massive Bio Call Center:
844-627-7246
Email: Support@massivebio.com

Facesheet

First Name: Jane	Classification:
Last Name: Doe	Histology Type: Squamous cell carcinoma, NOS
Birthdate: 1/23/1973	Country: United States
Age: 48	State: New Jersey
Sex: Female	City: Cliffside Park
Cancer Type: Skin	Zip: 07010
Stage: 4	Medical Oncologist: Dr. Carvajal

Comprehensive Summary

We have identified top 2 recommendations for you.

- NCT03844941: Phase II Randomized Trial of Avelumab Plus Cetuximab Versus Avelumab Alone in Advanced Cutaneous Squamous Cell Carcinoma of the Skin (cSCC). (Exact)
- NCT03822117: A Phase 2, Open-Label, Single-Arm, Multicenter Study to Evaluate the Efficacy and Safety of Pemigatinib in Participants With Previously Treated Locally Advanced/Metastatic or Surgically Unresectable Solid Tumor Malignancies Harboring Activating FGFR Mutations or Translocations (FIGHT-207) (Partial)

Please review the report. We can have a phone call with you to further narrow down the recommendations and fully discuss the identified options at your convenience.

Advances in application programming interfaces (APIs) and plug-in technologies will make it easier to achieve a relatively seamless interface between EHRs and genAI applications developed by third parties.

Massive Bio's Capabilities

Revolutionizing Oncology Value Chain



Clinical Trial Enrollment

Patient Recruitment



Patient Identification

Leveraging direct and indirect acquisition channels to identify eligible patients.



Pre-screening

AI-driven analysis of medical records and diagnostics to streamline pre-screening.



Clinical Trial Matching

Harnessing AI to precisely match patients with 19K+ actively recruiting clinical trials



Concierge Enrollment Enablement

Optimizing site selection and addressing logistical and financial barriers..

Beyond Patient Recruitment



NGS & Biomarker Trial Optimization

Assay selection, workflow design, and patient pre-screening for biomarker-based trials.



Sponsor Engagement & Data Integration

Acting as a Patient Contact Center for sponsors, integrating commercial data.



Diversity & Inclusion in Clinical Trials:

Analyzing patient demographics to improve inclusivity in clinical research.



Advanced Site & Trial Support

Validating sites, integrating reimbursement vendors, managing patient outcomes data.



Data and Analytics



Timely Alerts & Insights

Timely triggered alerts to surface treating physicians, enabling timely sales rep. visits.



Physician Profiling

Identifying behavioral trends to optimize execution strategy through physician profiling.



AI Driven Precision

AI-driven matching optimizes commercial spend with precise medical record analysis.

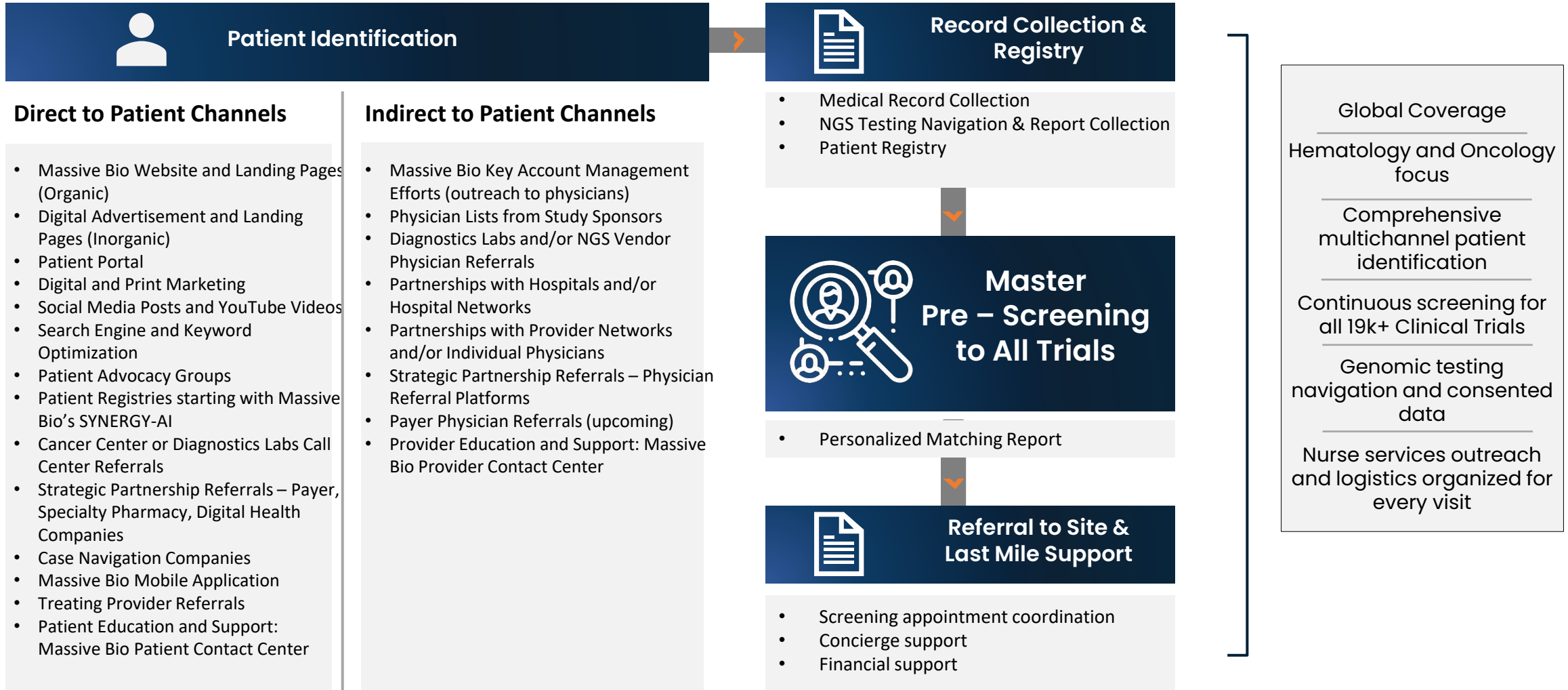


Value for Patients

Advanced analytics and genomics identify patients, enabling personalized care

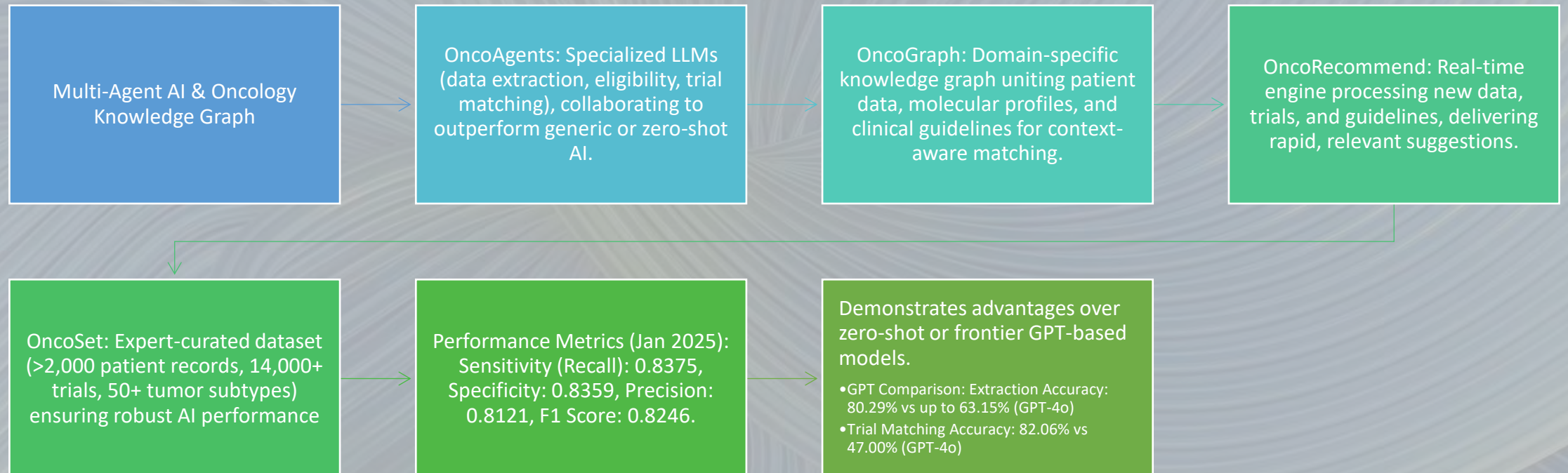
Disrupting Clinical Trial Enrollment

Master Prescreening Hubs



SYNERGY-AI

AI platform that automates medical data extraction, leverages an oncology-specific knowledge graph, and provides real-time trial recommendations, demonstrating a significant reduction in staff effort while maintaining high clinical accuracy.



Diversity Activation Model

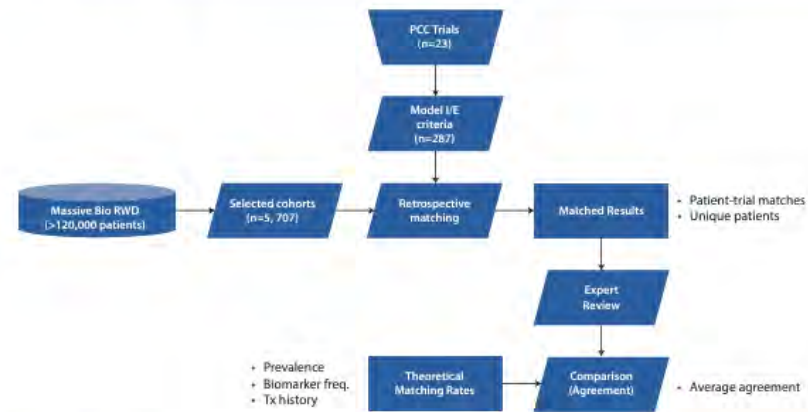
Massive Bio's recruitment techniques are, by nature, engineered for recruiting diverse populations. We search for patients in small community practices where there are dense patient populations which boosts diversity. Below are a few of current diversity initiatives:

- **Direct to Patient Diversity Baseline**
 - Race and ethnicity data is collected through patient's consent form, calls with patients, and medical records.
 - US patient acquisition is focused on community oncology.
 - Efforts outside of the US take place in over 17 countries.
- **Diversity Related Initiatives**
 - Leveraging proprietary registry, EMR, and claims data to target minorities and community-based practices serving those with unmet needs.
 - Intensifying focus on patient advocacy partnerships with groups who serve minority groups or those with unmet needs.
 - Increasing efforts to partner with physicians and facilities who serve minority groups or work in low-income areas.
 - Developing educational content demonstrating the importance of minorities taking part in clinical research to build trust and confidence for those seeking care.
- **Enterprise Level Initiatives Promoting Diversity**
 - In 2023, Massive Bio launched the "I Have Cancer" campaign. This is a highly people-focused marketing campaign which leverages diversity in the marketing imagery and explores narratives of individuals from around the world who have or are currently experiencing cancer in their life.



Study design

- The unique AI system extracted 180 structured clinical parameters from the patients' medical records and used a decision-tree algorithm to retrospectively match them to digitized inclusion/exclusion (I/E) criteria from over 14,000 actively recruiting interventional cancer trials. With particular focus on tumor types relevant to 23 selected trials.
- Results were compared to the theoretical matching rate based on specific criteria including tumor type, biomarker prevalence, disease extent at diagnosis, and prior treatment history.



The PCC is a non-profit organization dedicated to driving global access to comprehensive genomic testing for all patients with cancer. The PCC is composed of and funded by: AstraZeneca, Bayer, GSK, Johnson & Johnson, Eli Lilly, Novartis, and Roche, at the time of writing.

PCC-061 Date of preparation: June 2024



#ASCO24 Abstract e13501

Improved trial-matching outcomes A Massive Bio-PCC collaborative project demonstrates excellent outcomes

The combination of **Next-Generation Sequencing (NGS)** with **Artificial Intelligence (AI)** in a multi-trial matching approach resulted in:

Nearly 2-fold increase in potential patient-eligibility for trials for all tumor types

2-fold increase in a single focused tumor profile

Total saving in effort was estimated to be 19,500 hours (99.9%)

Introduction and purpose

- Patients' trial options and overall enrollment rates are severely restricted as a result of manual patient-screening by clinical research teams.
- To address this issue, the Precision Cancer Consortium (PCC) collaborated with Massive Bio to explore an innovative and efficient multi-trial matching method.
- They developed a multi-trial prescreening approach, incorporating NGS testing results, and using artificial intelligence.

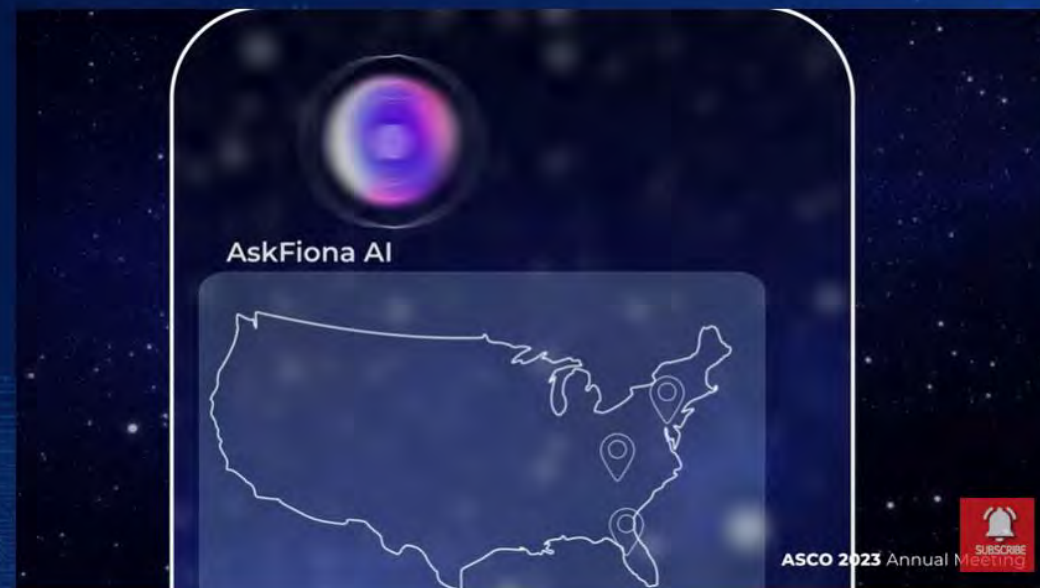
Results

- Accuracy and efficiency of trial matching was enhanced, particularly for targeted therapies.
- The total number of matching pairs (patient-trial) was 1,254 resulting in a **1.82-fold increase** as a result of multi-trial matching.
- Compared to the theoretical matching rate, **Massive Bio's matching rate was significantly higher** 50% of trials, and equivalent in 30%.
- The total savings in effort was estimated to be 19,500 hours (99.9%).
- Read more about the lung cancer case study in patients targeted for EGFRm and the reported 12-fold increase in patients on the abstract (scan the QR code).



DrArturo AI Augmented Intelligence Chatbot

DrArturo AI comes into play when a patient is deemed eligible for more specific information about clinical trials. He offers detailed insights about individual trials, as if he were a physician explaining the intricacies to a patient.





Massive Bio Clinical Network™

Welcome

Please enter your details.

Email

Password

☐ Remember for 30 days

[Forgot Password?](#)

Sign In

Sign in with

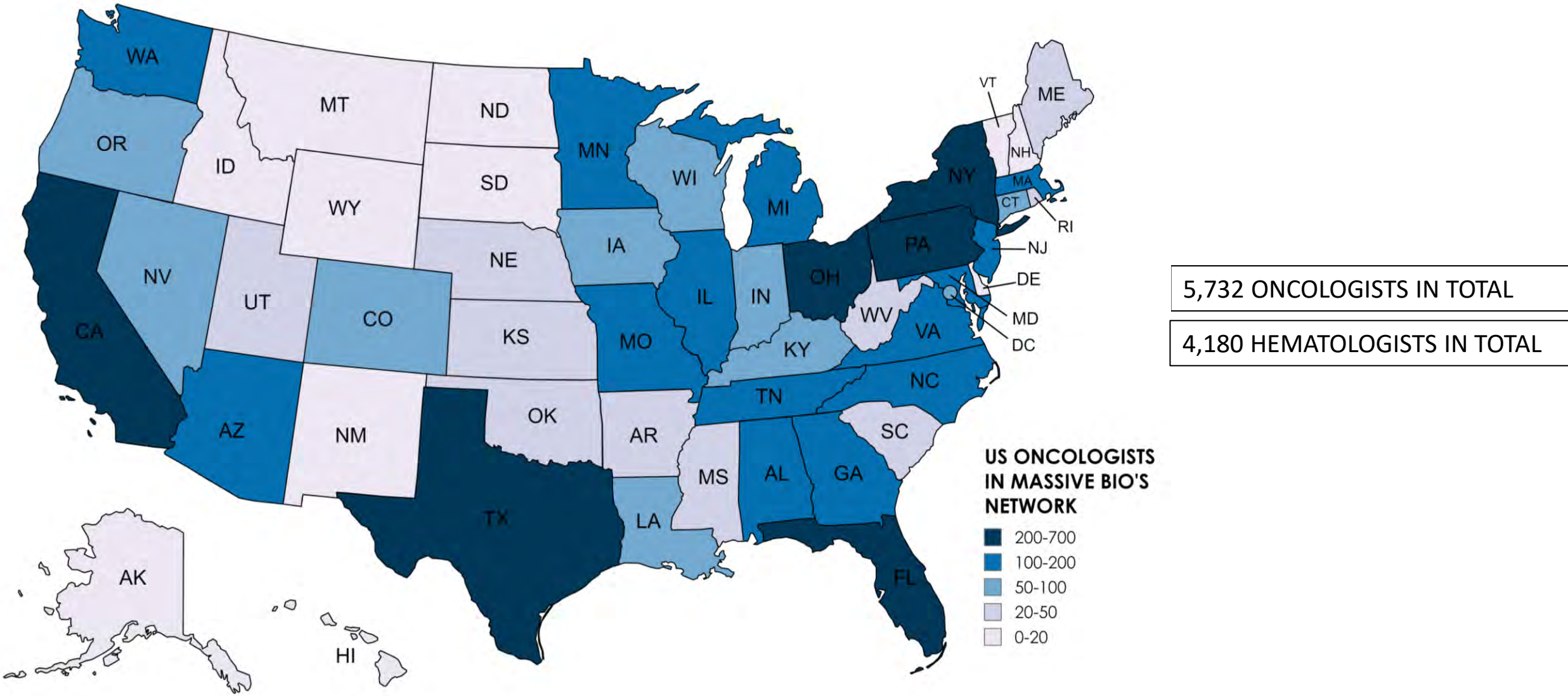


Don't have an account? [Sign up](#)

© Massive Bio 2024



US Oncologists in Massive Bio's Network



Solving the Paradox: Bringing the science closer to every patient

Leveraging AI, Patient Centricity, Precision Oncology and Just in Time Networks for Clinical Trial Enrollment

Reduce burden on participants to improve access. A large part of that is incorporating elements of a decentralized clinical trial, such as remote technology and measurement.

Modern innovations—from artificial intelligence (AI) to digital tools and analytics software—can help us accelerate the progress we’ve already begun to make and truly move the needle on health equity.

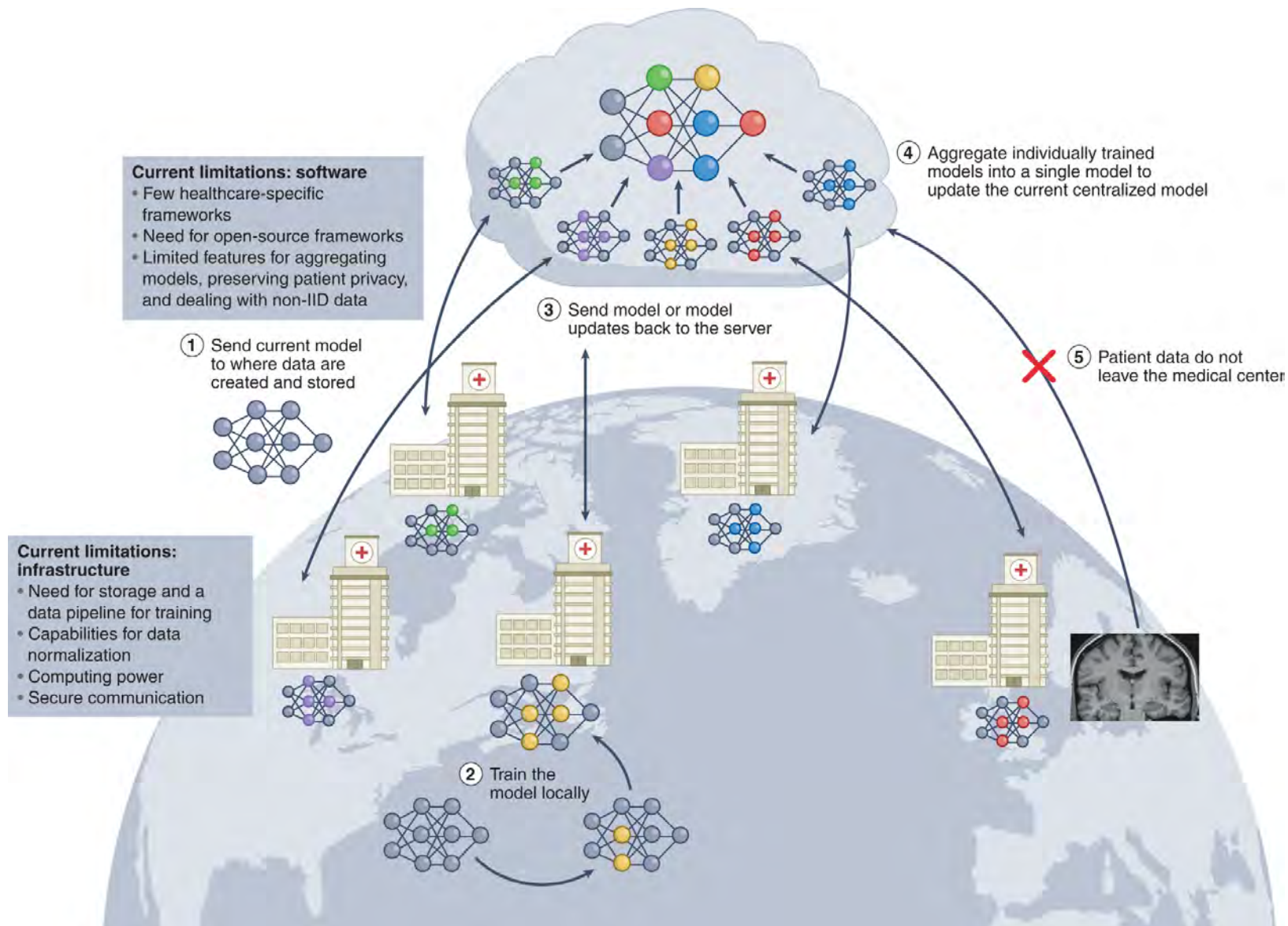
Decentralized, just-in-time, and hybrid clinical trials are already helping to alleviate the issues of location, cost, and time.

These trials proving to be more efficient than site-based trials, but they are also more inclusive.



Decentralized Clinical Trials



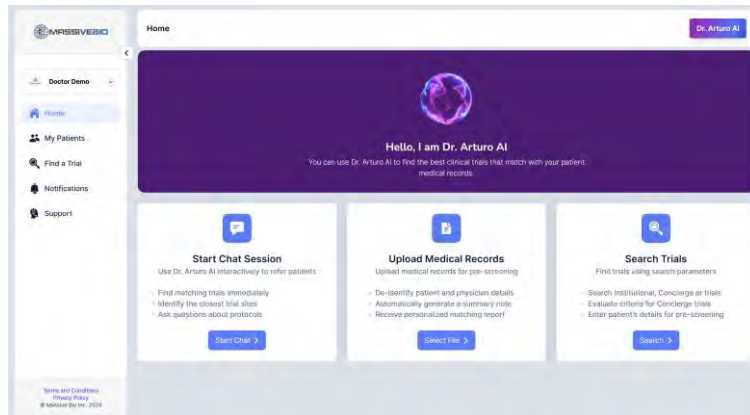


Clinical Network

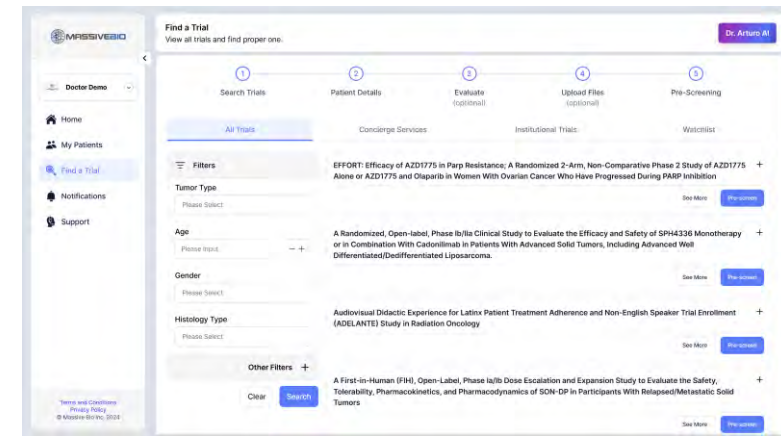
Clinical Network (CN) simplifies the process for physicians to refer patients to suitable clinical trials, leveraging AI-powered software and Massive Bio's concierge services. [For CN Video Click Here](#)

- Up-to-date information on actively recruiting clinical trials.
- Ability to evaluate eligibility criteria for multiple trials.
- Referring efficiently using speech or text multilingually.
- Systematic pre-screening and matching process.
- Medical record deidentification with HIPAA, GDPR and PIPEDA.
- Comprehensive database with detailed patient population.

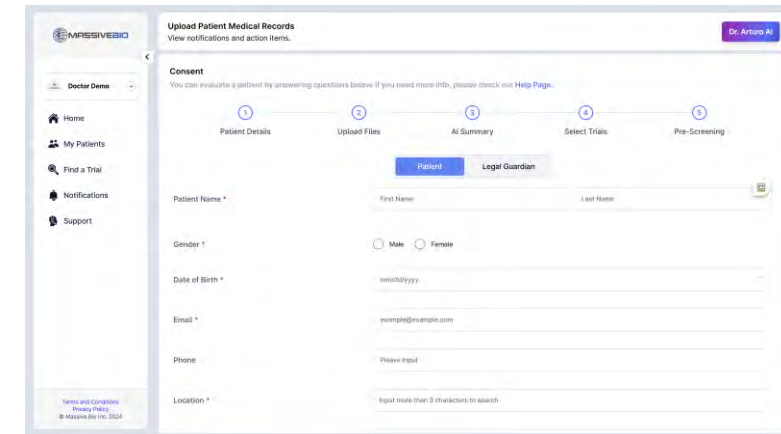
One-stop Physician Hub: AI Chat, Medical Record, Referral and Search



Global Trial Information: Find a Trial and Refer a Patient to Concierge Trials



AI Capabilities: Refer a Patient via AI supported Medical Record deidentification



Massive Bio Recognized by White House Cancer Moonshot for First-of-Their-Kind Pre-Screening Hubs and AI-Powered Trial Matching, Launches \$15 Million Initiative at JPM 2025

WH.GOV  

JANUARY 15, 2025

FACT SHEET: Marking Historic Progress, the Biden Cancer Moonshot Convenes Mission Report and Announces New Government and Private Sector Actions to Accelerate Progress Against Cancer


BRIEFING ROOM
STATEMENTS AND RELEASES

Nearly three years after the reignition of the Biden Cancer Moonshot, the work to end cancer as we know it has had a direct impact on American lives.

WH.GOV  

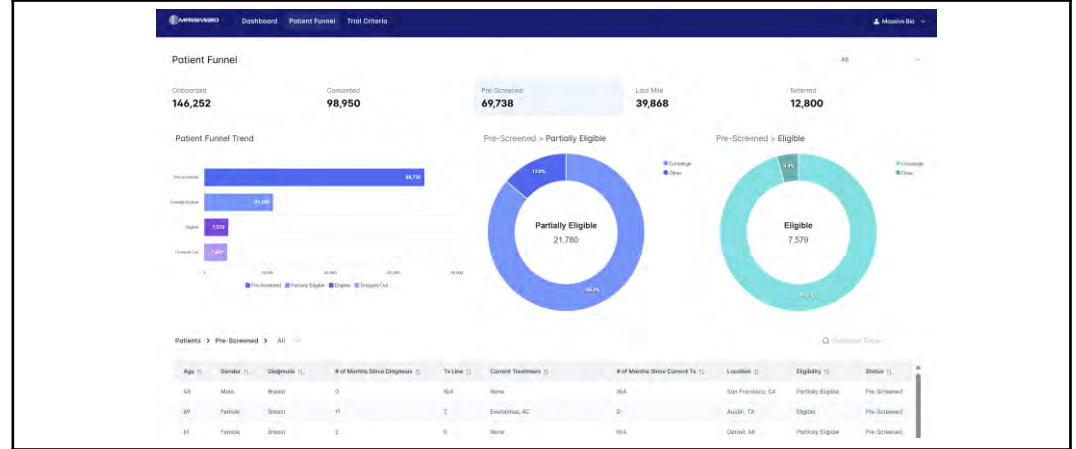
Massive Bio is launching an advanced AI-powered clinical trial matching platform and pre-screening hubs to empower patients and communities with culturally competent, seamless trial navigation and ongoing support. This system ensures timely, personalized clinical trial access by leveraging real-world data and comprehensive genomic profiling to match patients with optimal trial opportunities. Launching in January 2025, this expansion initiative represents an initial investment of \$15 million, and is expected to impact over 50,000 unique patients living with cancer annually, with a scalable goal of reaching 250,000 patients per year globally by 2027. The platform particularly focuses on underserved and underrepresented populations to ensure equitable access to clinical trials. Through this effort, Massive Bio is bridging critical gaps in care, accelerating access to the latest cancer treatments and improving outcomes for all.

Trial Vision

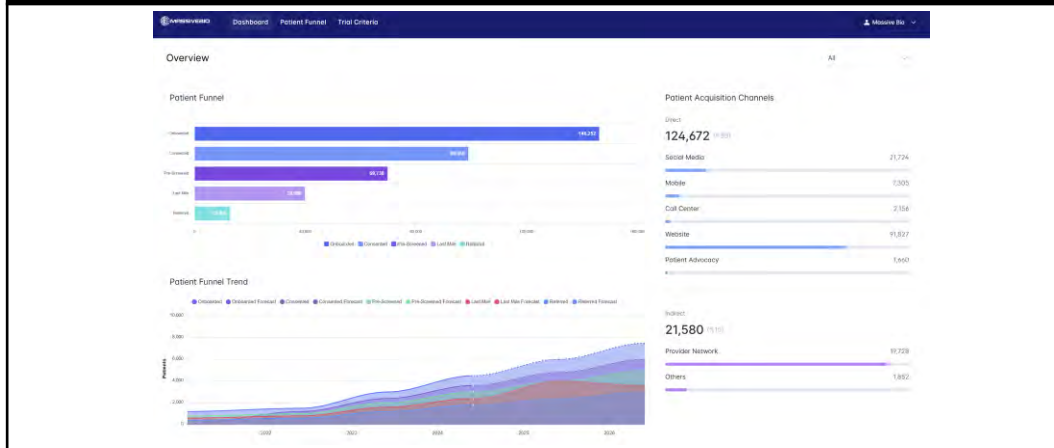
Trial Vision is an innovative platform designed to empower sponsors and clinical operations teams with comprehensive oversight of clinical trial progress. [For Trial Vision Video Click Here](#)

- Real-time visibility into the status and milestones of ongoing clinical trials.
- Insightful analytics on the funnel dynamics, highlighting the success of inclusion and exclusion criteria for trials.
- Updates on referred, eligible or ineligible for referral patients.
- Trial status comparisons.

Patients Eligibility: Exact & Partial Matches



Acquisition Insights: Patient Funnel, Trends & Channels



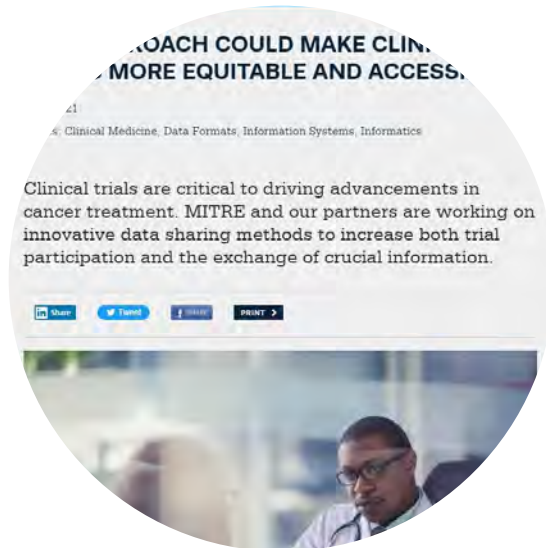
Trial Criteria: Matching Status & Distribution



Power through collaboration to maximize societal impact

















CO-HOSTED BY



Scale of Patient Acquisition

Our alliances drive Massive Bio's potential, positioning us for **substantial growth** and **impact**

 	1.8 million patients 100+ clinicians 70+ clinical locations
 	150,000 patients
 	6 million patients 250+ locations
 	1.3 million patients 8,578 pharmacies
 	600,000 patients 7,400 patient access points
 	2.8 million patients 60,000 clinicians 2,000 clinical locations
 	500,000+ patients' genetic information

 **4 million patient records**
100+ clinical locations

 **500,000 patients**

 **500,000 genomic profile analyses**

2024

160K+ Patients Onboarded

3.4K+ Sites (2.5K recruiting, 0.9K not yet recruiting)

100K+ Physicians

2025 & Beyond

19M+ Patients & Patient Data*

4K+ Sites

160K+ Physicians

**The total patient count indicates the number of access points to cancer patients and may include duplications from overlapping patient records in multiple partnerships.*

Unparalleled patient & provider scale

 **Hospital System**  **Payer**  **Patient Advocacy**  **Pharmacy**  **Laboratory**  **Data Company**

Massive Bio's Universe of Global Partnerships

Strategic Partners by Type	
Partner Type	Number of Partners
Hospital or Hospital Network	31
Digital Marketing Company	19
Provider Network	19
CRO	16
Digital Health Company	10
Insurance Company	8
Data Company	4
Patient Recruitment Company	3
Consortium	2
Pharmacy	2
SMO	1
Grand Total	131

Massive Bio and Foundation Medicine Partner to Optimize Clinical Trial Recruitment and Enrollment to Expand Access to Precision Cancer Care

September 13, 2024 | 4 min read



NEW YORK--(BUSINESS WIRE)--Massive Bio, a cutting-edge AI driven clinical trial matching platform provider, and Foundation Medicine, Inc. today announced a strategic collaboration to increase patient access to cutting edge clinical trials. This collaboration unites Massive Bio's patient matching technology and last-mile solutions with Foundation Medicine's global expertise and leadership in next generation sequencing (NGS) and network of healthcare providers.

Massive Bio partners with Foundation Medicine to enhance clinical trial recruitment and expand access to precision cancer care.



Walgreens

Quest Diagnostics

Optum

vithas

Tzafon Medical Center

RARE PATIENT Voice

naic

HOSPITAL SAN AGUSTÍN

PRECISION CANCER CONSORTIUM

NEO GENOMICS

EVERNORTH HEALTH SERVICES

BlueCross BlueShield

quirónsalud

IMOMA

SZPITAL UNIWERSYTECKI

MPN

ISTITUT ROMANEO PER LO STUDIO DEI TUMORI DINI AMADORI

ONCOASSIST

CANCER TREATMENT CENTERS OF AMERICA

DATAVANT

trialbee

HOSPITAL UNIVERSITARIO CLÍNICO SAN CECILIO

The Oncology Institute of Hope & Innovation

trinal

RAKIETY

euofins

further

healthincode

Perthera

CUREMATCH

CANCERX

Patient Advocacy Group Engagement

Number of Patient Advocacy Groups by Country	
Country	Number of Groups
United States	230
Spain	77
France	20
Greece	20
Italy	20
Germany	15
Brazil	10
Belgium	4
Israel	4
Romania	4
Poland	3
Switzerland	3
Turkey	3
United Kingdom	3
Grand Total	416

- There are two primary ways that Massive Bio identifies patient advocacy groups:
 - We identify groups independently from a specific sponsor or protocol and work with the group for all relevant trials
 - The sponsor introduces Massive Bio to a advocacy group and Massive Bio works with the group to establish linked marketing and support in other ways
- Massive Bio has an established relationship with 416 advocacy groups.
- As part of the onboarding process for a new trial, Massive Bio will target and onboard new advocacy groups.
- Groups are either cancer-specific and support patients with only certain cancer types or they are more general and will support patients with any type of cancer.
- For both online-only groups and brick & mortar organizations, Massive Bio will assign a point of contact in the specific country and typically outreach to the group a minimum of once a month.

Significant Understanding of NGS Testing & Relationships with Vendors that Minimizes Testing & Screening Costs

NGS Vendor Breakdown of Patients who Come to Massive Bio's Patient Contact Center	
NGS Provider	Percentage
Foundation Medicine	26%
Caris	17%
Large Cancer Center Lab/Other	17%
Tempus	11%
Guardant	10%
Invitae	5%
NeoGenomics	3%
Myriad	3%
Ambry Genetics	2%
MD Anderson Cancer Center	2%
Memorial Sloan Kettering Cancer Center	2%
Strata	1%
Quest	1%
Genpath	<1%
PGDX	<1%
Grand Total	100%

- Massive Bio facilitates the NGS testing workflow for patients that come through our contact centers in cases where the patient has not been tested. We support the physician by assisting with the completion of NGS testing request forms and delivery of the request to the NGS vendor. As part of medical records collection, we obtain the NGS test results which we can then send to the sponsor in a de-identified format with the patient's consent.
- Massive Bio supports NGS vendors by providing last mile services to the physicians that have patients with the required biomarker. For example, if the physician doesn't have time, Massive Bio onboard the patient at its patient contact center, prescreens and send the prescreening results to the treating oncologist.
- Massive Bio analyzes eligible patient volume for potential just-in-time sites before activation by an NGS vendor such as Caris or Tempus.

Massive Bio's Differentiators

Creating a scalable, tech-enabled channel to accelerate and enhance the entire oncology value chain



Global Presence & Reach

- Operating in **17 countries** with 84 employees across the United States, Canada, Europe, LATAM, Israel, and Turkey
- Proprietary AI system scans over **19,000 clinical trials** worldwide to match patients with eligible trials



Strategic Partnerships

- Partnership with **131 strategic** organizations, including Foundation Medicine, Flatiron, Walgreens, Quest, and Optum
- Engaged with **416 patient advocacy groups** worldwide, ensuring a broad and inclusive patient outreach



Oncology & Hematology Expertise

- Specializes in oncology and malignant hematology, having been hired for **127 oncology/hematology studies** since 2019
 - **Biomarker**-based recruitment for precision oncology, guided by genomic data and relationships with key NGS vendors



Patient Consented Data

- Utilizes **patient consented data** to ensure compliance and to enhance recruitment and tracking processes across various platforms
- **Direct & indirect patient identifications** channels ensuring extensive outreach



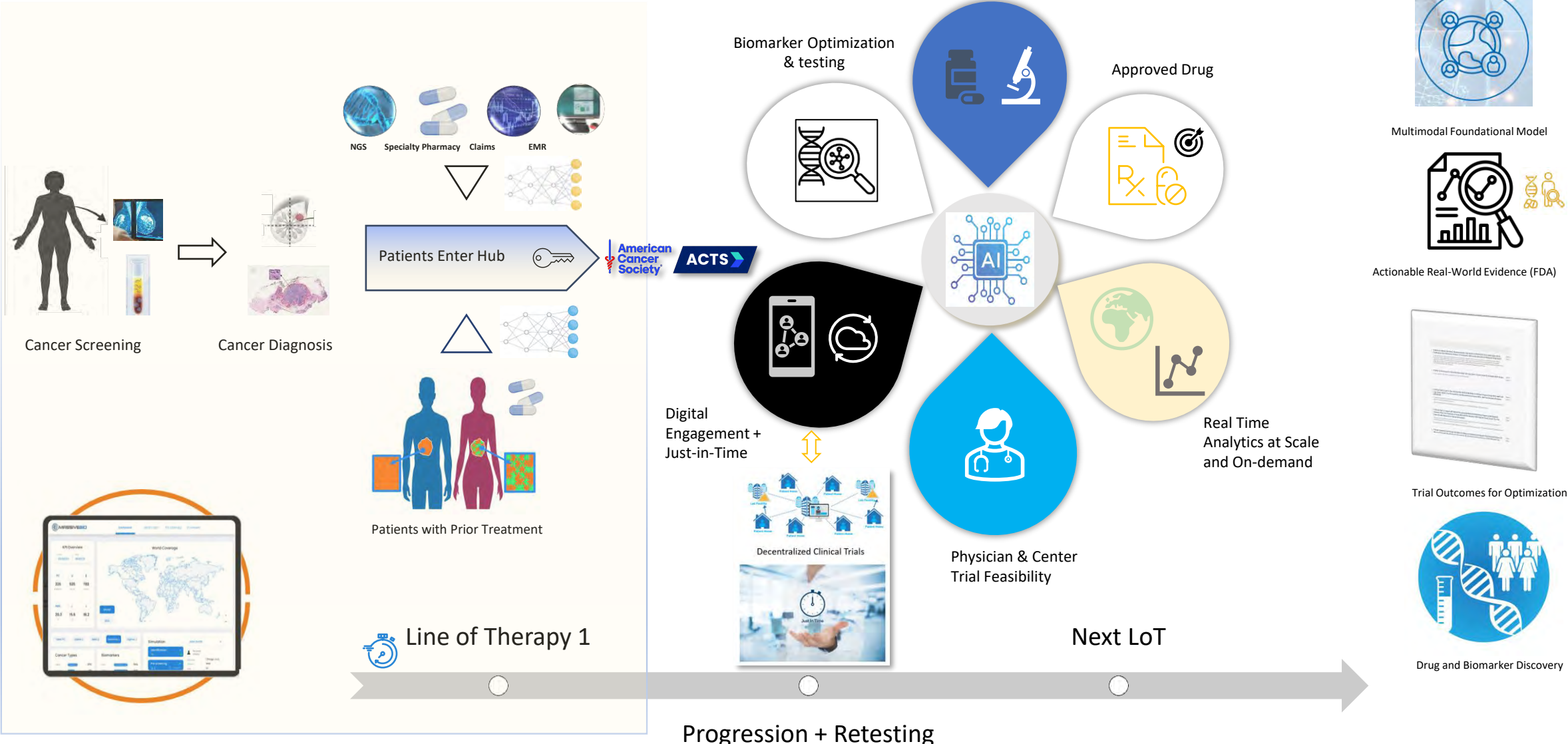
Last Mile & Concierge Services

- **Last-mile** patient guidance, ensuring successful enrollment and follow-through in clinical trials
- Nurse services and **logistics**, ensuring comprehensive support from initial consultation to trial participation.

Pre-screening hub-and-spoke population wide approach

Patient centric just-in-time, decentralized/hybrid trials

Transforming every patient's journey into a catalyst for discovery





**Clinical Trials
Education**



**Health-related
Social Needs
Screening**



**Trial Eligibility
& Enrollment
Navigation**



**AI-powered
Clinical Trials
Matching**



**Transportation
Assistance
to Trials**



**Lodging
Support Near
Trial Sites**

Paradox Solved: The Future of Clinical Trial Matching



As we begin a new quarter, Massive Bio continues its global efforts to enhance clinical trial access through AI-driven solutions, strong physician engagement, and strategic partnerships. In this edition, we're sharing recent recognition, highlights from international activities, and upcoming events where we hope to connect with you.

Highlights from This Edition

- News and Announcements
- Meet Us at These Events
- Global Provider Engagement Updates
- Clinical Network Update

News and Announcements

ACS and Massive Bio Working Together to Expand Clinical Trial Access

The American Cancer Society and Massive Bio are working together to make it easier than ever for patients to find and access cancer clinical trials. This collaboration is focused on improving awareness and access—empowering more people to explore potential options through clinical research. You can [click here](#) to explore partnership opportunities.



Access Clinical Trials With **ACS ACTS™**

Finding the right clinical trial can be tough. ACS ACTS offers free, personalized clinical trial matching and resources. Available in select states.



Scan the code, call us at 1-800-227-2345, or visit [cancer.org/acts](#) to learn more!



[cancer.org/acts](#)



Access Clinical Trials With **ACS ACTS™**

Finding the right clinical trial can be tough. ACS ACTS offers free, personalized clinical trial matching and resources. Available in select states.



Scan the code, call us at 1-800-227-2345, or visit [cancer.org/acts](#) to learn more!



[cancer.org/acts](#)

It is up to US to use the potential of AI to the fullest in the next decade and remove all possible barriers





Théâtre D'opéra Spatial

PERSPECTIVE

Harnessing Moravec's Paradox in Health Care: A New Era of Collaborative Intelligence

Arturo Loaiza-Bonilla , M.D., M.S.Ed., F.A.C.P.^{1,2} and Scott Penberthy , Ph.D.³

Received: January 3, 2025; Revised: January 22, 2025; Accepted: February 12, 2025; Published: April 17, 2025

Abstract

Artificial intelligence excels at complex data analytics, yet struggles with nuanced, sensorimotor tasks that humans perform almost effortlessly — a dichotomy encapsulated by Moravec's paradox. By strategically harnessing these complementary strengths, health care can usher in an era of collaborative intelligence, optimizing data-intensive workflows, such as clinical trial enrollment, and creating more patient-centric models of care.

Introduction

Moravec's paradox highlights a counterintuitive truth: advanced mathematical or data-driven tasks are often easier for computers than the everyday sensorimotor and social tasks humans perform and take for granted.¹ Nowhere is this more evident than in health care. A seasoned clinician can, within seconds, detect a patient's anxiety or discomfort by reading subtle body language — yet manually sifting through thousands of laboratory results or complex clinical trial protocols can be laborious and time-consuming. Meanwhile, artificial intelligence (AI) systems can handle those data-heavy tasks at superhuman speed, but remain clumsy with seemingly easy tasks like empathetic bedside manner or dexterous physical procedures.

This tension need not be a limitation. By aligning each domain — human and machine — with the tasks to which it is best suited, we can create a collaborative model of care that increases efficiency and, more importantly, preserves a deeply humane approach to medicine.

The Evolutionary Clue

Why is this mismatch so prominent? Evolution devoted vast energy to refining our sensorimotor and social skills — abilities critical for survival and social bonding over millions of years.¹ By comparison, advanced formal reasoning and numerical computation are relatively recent additions to our cognitive repertoire. AI research, conversely, has historically focused on algorithms for large-scale computations and pattern recognition. It is thus unsurprising that algorithms excel at high-dimensional data analysis well before mastering the physical dexterity or empathic presence we expect of a clinician, and concepts of artificial general

The author affiliations are listed at the end of the article.

intelligence and artificial superintelligence are emerging,² including the assumption of medical workforce shifts.³

Easy for AI, Hard for Humans: Advanced Data Analysis and Prediction

COMPLEX DATA SYNTHESIS

AI-based models can integrate massive datasets — including electronic health records (EHRs), imaging, and genomic information — and detect patterns beyond the scope of human specialists.³ This expands opportunities to stratify risk, forecast patient outcomes, and discover subtle correlations in diverse patient populations. From a technical standpoint, these capabilities rely on advanced machine learning (ML) pipelines that typically include data cleaning, normalization, and feature engineering stages before modeling. In many modern workflows, natural language processing (NLP) tools extract key clinical entities (e.g., medications, diagnoses, biomarkers) from unstructured notes. These structured data are then aggregated with imaging and genomic inputs for multimodal analyses using techniques such as deep neural networks, gradient-boosting machines, or transformer-based architectures. Each step is validated using rigorous cross-validation protocols and external holdout sets to ensure generalizability. Additionally, attention-based mechanisms within certain models (e.g., vision transformers for imaging⁴ or large language models for text⁵) help highlight the most salient features driving predictions. This interpretability is crucial for clinical adoption, as it allows clinicians to understand why an AI system flags certain data points as anomalous or high risk, thereby fostering trust and transparency in AI-driven insights.

AUTOMATING REPETITIVE BUT HIGHLY COGNITIVE TASKS

Billing, coding, protocol optimization, or summarizing large blocks of text data can drain a clinician's time and attention. ML systems handle these tasks rapidly, freeing professionals to focus on higher-level clinical decisions. These automations often leverage NLP models — ranging from more traditional approaches (like conditional random fields or rule-based systems) to advanced generative pretrained transformers (generative pretrained transformer-like architectures). By training on large corpora of clinical notes or administrative data, they can learn domain-specific terminology and coding structures. The result is expedited processing of claims

or protocol documentation with lower error rates compared with manual approaches. However, ongoing monitoring of model outputs is essential to detect concept drift — where shifting medical standards or coding guidelines may degrade performance over time. A robust quality-assurance loop, led by clinical domain experts, ensures that any automated recommendations align with current best practices and maintain regulatory compliance.^{6,7}

Hard for AI, Easy for Humans: Empathy, Dexterity, and Contextual Flexibility

BEDSIDE COMMUNICATION

Humans excel at sensing microexpressions or shifts in voice tone, and tailoring responses accordingly. While chatbots may mimic empathy textually, they lack the deep emotional resonance and adaptability of human clinicians.

PHYSICAL DEXTERITY

Basic procedures like drawing blood or palpating an organ during a physical exam, and any surgical procedure, will require sensorimotor finesse that eludes contemporary robotics. The simple manual skills that health professionals refine over years remain a frontier challenge for AI.

CONTEXTUAL FLEXIBILITY

Physicians often adjust care by accounting for nonverbal patient cues, cultural backgrounds, and socioeconomic constraints.⁸ Such holistic reasoning is deeply human, and not readily captured by current AI models.

A Hybrid Future in Clinical Work

Modern medicine encompasses both the physical (e.g., exams, surgeries) and the abstract (e.g., analyzing large and complex datasets, managing elaborate protocols). Collaboration between clinicians and AI can address these distinct domains, with clinicians providing real-time empathy, creativity, moral judgment, and problem-solving in unpredictable settings, while AI rapidly mines large datasets, flags anomalies, and proposes evidence-based pathways drawn from extensive clinical knowledge.

This synergy enhances efficiency without compromising the personal touch. Rather than pursuing a fully autonomous medical robot, a more practical approach is a

human-machine partnership that offloads tedious computational tasks from clinicians, allowing them to focus more on relational and hands-on care.

Decentralizing Clinical Trials with AI as a Case Study

A powerful demonstration of Moravec’s paradox in action is the use of AI in clinical research⁸ (what we call the solution to the clinical trial enrollment paradox).

- **Rapid Eligibility Screening:** Trial enrollment often involves scanning reams of inclusion/exclusion criteria. By automating this advanced data matching, AI drastically shortens recruitment timelines.
- **Decentralized Patient Recruitment:** Leveraging AI to handle data-intensive steps allows studies to expand beyond large academic centers, increasing access for diverse patient populations.
- **Balancing Promise with Practicality:** Challenges persist — algorithmic biases, privacy concerns, and the need for validation remain critical. Data analysis can be automated, but humans must still guide ethical oversight and informed consent processes.

Having cofounded an AI-driven initiative focused on clinical trial matching, and from having experience in working at the intersection of AI and biomedical sciences, we have witnessed firsthand how quickly data-heavy tasks can be automated, with physicians still playing an indispensable role in communicating trial options, providing clarifications, and shaping patient-centric approaches.⁹

Reflecting Moravec’s Paradox: Tasks Likely to Be Automated First

Given Moravec’s insight that formally reasoned, data-driven tasks come more naturally to machines than sensorimotor tasks (Fig. 1), these tasks present opportunities for automation to enhance clinical efficiency, decision support, and patient care.

- **EHR Data Summaries:** Automated systems can consolidate patient histories into concise profiles for quick physician review.
- **Protocol and Clinical Guideline Optimization:** AI can suggest trial modifications or guideline updates based on emerging data.

Easy for Humans (Hard for AI)	Easy for AI (Hard for Humans)
Empathy and bedside communication	Large-scale data analysis
Physical dexterity in exams	Rapid trial eligibility scanning
Contextual flexibility	Automated EHR summarization

Figure 1. Moravec’s Paradox in Health Care.

AI denotes artificial intelligence; and EHR, electronic health record.

- **Remote Triage and Screening:** Online questionnaires can be processed by AI to determine preliminary risk, reserving direct human attention for critical cases.

These data-driven tasks often hinge on sophisticated pattern-recognition algorithms. For instance, EHR data summaries can be automated through hierarchical deep learning models that sequentially parse structured and unstructured data (e.g., laboratory values, imaging reports, clinician notes) to craft concise, contextually relevant clinical overviews. Similarly, AI-assisted protocol and guideline optimization can leverage reinforcement learning techniques to iteratively test modifications in virtual environments — simulating how changes to inclusion criteria or dosing regimens might affect outcomes. However, translating these insights into real-world clinical workflows requires robust, reproducible validation methods. Studies should follow frameworks for evaluating clinical prediction models akin to PROBAST (Prediction model Risk Of Bias ASsessment Tool)¹⁰ or TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis)¹¹ to ensure methodological rigor in model development and reporting, thereby safeguarding against common pitfalls like overfitting, selection bias, or hidden confounders.

Meanwhile, high-dexterity procedures, dynamic bedside communication, and emergent problem-solving that hinge on empathy or creativity remain firmly in the human domain.

Addressing Algorithmic Bias, Validation, and Real-World Implementation

Although AI has the potential to accelerate clinical research and improve care delivery, it also risks perpetuating or even amplifying existing biases. Algorithmic bias can emerge

if training data are not representative of all patient demographics, leading to disparities in diagnostic accuracy or treatment recommendations.¹² To mitigate this, developers must implement strategies such as diverse data curation, bias audits, and fairness metrics during model building and deployment.

In parallel, model validation must go beyond traditional internal validation; external validation on multi-institutional datasets is key to ensure generalizability across populations with varying genetic, environmental, and socioeconomic backgrounds. Prospective clinical trials or real-world evidence registries can then confirm the effectiveness of these AI-driven interventions in practice.

Finally, for models that are successfully validated, attention must be paid to ethical and regulatory requirements. Institutions may convene multidisciplinary oversight committees that include clinicians, data scientists, patient advocates, and ethicists to establish frameworks for responsible AI adoption. Such committees can evaluate real-time model performance, respond to unexpected behaviors, and adjust governance measures as needed, ensuring that AI becomes a trusted, transparent partner in health care rather than a black box solution.¹³

Conclusion

Moravec's paradox points toward a future of collaborative intelligence rather than a battle for domain supremacy between humans and machines. Advanced analytics, powered by AI, can transform data-rich processes in health care — from triaging and risk scoring to matching patients with clinical trials. Yet, it is the human clinician's capacity for empathy, dexterity, and contextual insight that ensures patients receive truly holistic care.

This partnership, built on the complementary strengths of humans and AI, promises to lessen administrative burdens, promote more inclusive research participation, and ultimately improve patient outcomes. It also underscores a key principle: technology should serve as a force multiplier for compassion, not a replacement for it.

Disclosures

Author disclosures are available at ai.nejm.org.

Author Affiliations

¹Massive Bio, New York

²St. Luke's University Health Network/Lewis Katz School of Medicine at Temple University, Bethlehem, PA

³Office of the CTO, Google Cloud, Mountain View, CA

References

1. Moravec HP. Mind children: the future of robot and human intelligence. Cambridge, MA: Harvard University Press, 1988.
2. Klang E, Tessler I, Freeman R, Sorin V, Nadkarni GN. If machines exceed us: health care at an inflection point. *NEJM AI* 2024;1(10). DOI: [10.1056/AIp2400559](https://doi.org/10.1056/AIp2400559).
3. Eriksen AV, Möller S, Ryg J. Use of GPT-4 to diagnose complex clinical cases. *NEJM AI* 2023;1(1). DOI: [10.1056/AIp2300031](https://doi.org/10.1056/AIp2300031).
4. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: transformers for image recognition at scale. June 3, 2021 (<https://doi.org/10.48550/arXiv.2010.11929>). Preprint.
5. Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst* 2020;33:1877-1901. DOI: [10.48550/arXiv.2005.14165](https://doi.org/10.48550/arXiv.2005.14165).
6. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med* 2019;380:1347-1358. DOI: [10.1056/NEJMra1814259](https://doi.org/10.1056/NEJMra1814259).
7. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health* 2021;3:e745-e750. DOI: [10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9).
8. Kurnaz S, Loaiza-Bonilla A, Carvalho Castañeda D, Huner O, Giritlioglu D. Effect of a novel artificial intelligence (AI) - enabled multi-trial matching system on patient matching using real-world data. *J Clin Oncol* 2024;42:e13501. DOI: [10.1200/JCO.2024.42.16_suppl.e13501](https://doi.org/10.1200/JCO.2024.42.16_suppl.e13501).
9. Wan P, Huang Z, Tang W, et al. Outpatient reception via collaboration between nurses and a large language model: a randomized controlled trial. *Nat Med* 2024;30:2878-2885. DOI: [10.1038/s41591-024-03148-7](https://doi.org/10.1038/s41591-024-03148-7).
10. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1-W33. DOI: [10.7326/M18-1377](https://doi.org/10.7326/M18-1377).
11. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594. DOI: [10.1136/bmj.g7594](https://doi.org/10.1136/bmj.g7594).
12. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366:447-453. DOI: [10.1126/science.aax2342](https://doi.org/10.1126/science.aax2342).
13. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019;25:1337-1340. DOI: [10.1038/s41591-019-0548-6](https://doi.org/10.1038/s41591-019-0548-6).



Public Comment on the 2025 National AI R&D Strategic Plan RFI Docket ID: NSF-2025-OGC-0001

Subject: Response to Request for Information on the 2025 National Artificial Intelligence Research and Development Strategic Plan

(Submitted via Regulations.gov on April 30, 2025)

To the Networking and Information Technology Research and Development (NITRD) National Coordination Office (NCO) and the White House Office of Science and Technology Policy (OSTP):

Introduction

On behalf of [Massive Bio, Inc.](#), a leader in leveraging artificial intelligence to connect cancer patients to optimal treatments and clinical trials, thank you for the opportunity to provide input on the development of the 2025 National AI R&D Strategic Plan. As Co-Founder and Chief Medical Officer, I witness daily the transformative potential of AI to revolutionize healthcare, particularly in the complex field of oncology. We strongly support the Administration's commitment – as outlined in Executive Order 14179 – to sustain and enhance America's AI leadership in order to promote human flourishing, economic competitiveness, and national security ([Removing Barriers to American Leadership in Artificial Intelligence – The White House](#)).

Massive Bio's mission aligns directly with these goals. By utilizing [Synergy-AI](#) to analyze complex patient data (genomic, clinical, imaging) and match individuals to personalized therapies and cutting-edge clinical trials, we aim to improve patient outcomes, democratize access to advanced cancer care, accelerate drug development, and ultimately contribute to human flourishing. We believe that strategic federal investment in specific areas of AI R&D, particularly those with long-term horizons or complex societal implications like healthcare, is paramount to maintaining U.S. dominance and ensuring AI serves the public good. The previous National AI R&D Strategic Plan provided a valuable foundation. To secure America's position as the unrivaled world leader, the 2025 update must sharpen its focus on areas where federal leadership is indispensable – complementing rather than duplicating private sector efforts. This includes prioritizing foundational research, addressing systemic challenges like data access and bias, and fostering innovative ecosystems for collaboration. Healthcare, and specifically the fight against diseases like cancer, represents a domain where the confluence of high complexity, significant societal benefit, and market limitations necessitates robust federal R&D support.

Massive Bio itself is actively contributing to AI-driven healthcare innovation in ways that exemplify the kind of progress federal support can catalyze. We are strong supporters of the Advanced Research Projects Agency for Health's **Advancing Clinical Trials Readiness (ACTR)** program, developing AI-driven decentralized clinical trial platforms. Through these initiatives, we hope to deploy *federated* AI pipelines for privacy-preserving patient identification, multilingual electronic consent interfaces to broaden access, and longitudinal digital engagement tools to dramatically accelerate clinical trial enrollment and retention. Our goal – aligned with ARPA-H's – is to make 90% of Americans able to reach a clinical trial within 30 minutes of home, bringing trials to patients instead of requiring patients to travel. We have also spearheaded the development of large-scale, foundational AI models in precision oncology, leveraging *federated learning* across hospital networks so that sensitive data stays local while insights are shared. These efforts have demonstrated real-world impact: for example, integrating AI with genomic



testing doubled the identification of eligible trial patients and cut 19,500 hours of manual matching effort down to just 1 hour, showcasing how AI can exponentially scale clinical research.

Our experience underscores that visionary federal programs and public-private collaboration can enable breakthroughs that neither sector could achieve alone. With the right national R&D strategy, we can generalize such successes across the healthcare system. In this comment, we offer recommendations for bold yet practical federal action to advance AI in health and medicine, drawn from our on-the-ground insights and successes. We focus on priority research areas and enabling mechanisms that will ensure the United States leads in deploying AI to save lives, make healthcare more equitable, and drive economic and societal benefits.

I. Prioritized Federal AI R&D Areas for National Leadership (Next 3–5 Years)

While industry excels at applying existing AI techniques for near-term gains, federal investment is crucial for pioneering next-generation capabilities and tackling challenges aligned with the public interest – especially in healthcare. We recommend prioritizing R&D in the following areas:

Enhanced Patient Matching:

A critical innovation that must be federally prioritized is the deployment of digitally enabled AI prescreening hubs — centralized or distributed engines that continuously scan EHR systems for eligible clinical trial candidates. These hubs not only enable real-time trial matching but also provide a scalable framework to monitor data drift, ensuring that AI algorithms remain valid as clinical documentation practices, biomarker definitions, or population health trends evolve. Massive Bio has pioneered such architectures through its Synergy-AI platform, demonstrating how real-world EMR data can be transformed into structured eligibility insights at scale. Federally supported prescreening hubs can become the cornerstone for equitable access to research, making clinical trials a practical and visible use case for nationwide EMR interoperability. These systems would reduce physician burden, enhance trial awareness, and operationalize precision recruitment without violating patient privacy.

Interoperability and Data Standards:

The future of interoperable, patient-centric healthcare hinges on transforming clinical research into a flagship use case for real-time data integration across EMR platforms. Digitally enabled clinical trial matching systems — such as Massive Bio’s integration of AI, NLP, and federated learning — can serve as a proof-of-concept for this transformation. These platforms are already demonstrating how disparate EMRs, genomic reports, and imaging data can be harmonized in real time to identify life-saving opportunities for patients, while fully respecting HIPAA and patient consent frameworks. Clinical research thus becomes not just a consumer of data infrastructure, but its driver — pushing interoperability, patient-accessible tools, and consent governance into mainstream practice. A strategic plan that places clinical trials at the intersection of AI, interoperability, and patient empowerment can catalyze system-wide innovation.

Foundational AI for Complex Biological Systems

Need: Current AI excels at pattern recognition but struggles with the deep *causal reasoning* and multi-scale modeling required to understand the intricate dynamics of biology (e.g., tumor microenvironments, immune system interactions, drug resistance mechanisms). Industry focus

tends to remain on immediate predictive tasks, whereas decoding complex biological processes to enable breakthroughs in disease prevention or cure is a high-risk, long-term endeavor.

Priority: Invest in fundamental research on new AI architectures and algorithms (beyond current deep learning paradigms) capable of modeling complex, multi-modal biological processes. This includes AI that can integrate diverse data types (genomic, proteomic, imaging, clinical records) into holistic, dynamic models of human biology and disease. Such research should enable AI to not just identify correlations, but to infer causal relationships and mechanistic insights from biological data. This is a long-term challenge unlikely to receive adequate private funding initially, yet it is essential for breakthroughs in personalized medicine and drug discovery. Federal leadership can fill this gap by funding interdisciplinary teams (AI scientists, biologists, physicians) and computing infrastructure to develop these “digital twins” of biology. Early successes hint at the promise: for example, researchers recently developed an *explainable transformer* model that integrated multi-dimensional clinical and molecular data to predict cancer treatment outcomes, surpassing other methods and even identifying a new subset of patients who might benefit from immunotherapy. This kind of foundational AI model – which can generate hypotheses and insights beyond human intuition – illustrates the transformative potential of marrying advanced AI with complex biology. We must empower our research community to build the next generation of **AI-driven scientific discovery engines** that will accelerate understanding of diseases and the development of cures.

AI for Accelerated and Accessible to all Clinical Trials

Need: Clinical trials remain slow, expensive, and frequently struggle with patient recruitment and diversity. An estimated **80% of trials are delayed or closed due to insufficient enrollment** ([What clinical trial statistics tell us about the state of research today](#)), and *less than 5%* of adult cancer patients enroll in trials despite ~70% indicating willingness – a gap that leaves life-saving therapies untapped by most patients. Under-enrollment and lack of representation (e.g. of minority and rural populations) undermine both the speed and the generalizability of clinical research. While companies like Massive Bio and others are using AI to match patients to trials more efficiently, systemic challenges persist in trial design, site selection, real-time data monitoring, and integration of disparate health systems. We need AI to not just find the *right patient for the right trial*, but to fundamentally reinvent how trials are designed and conducted so they can be faster, more inclusive, and more effective.

Priority: Fund R&D into AI-driven tools and methodologies to modernize the clinical trial lifecycle, including:

- **Intelligent Trial Design:** AI models to optimize protocol criteria, simulate trial outcomes, and select endpoints in silico before a trial ever begins. By analyzing historical trial data and real-world patient populations, AI could help design studies that are more adaptive, efficient, and reflective of patient diversity.
- **Enhanced Patient Matching:** Advanced, privacy-preserving algorithms (e.g. federated learning and secure multi-party computation) to identify eligible patients across decentralized and diverse healthcare data systems, improving recruitment speed and representativeness. For example, federated machine learning can scan patient records at multiple hospitals without moving the data, achieving high accuracy in finding trial

candidates while preserving privacy. Such approaches can dramatically expand the pool of patients considered for trials – especially in community settings – and ensure that eligibility criteria are interpreted consistently (with AI reading unstructured notes, labs, genomics, etc.).

- **Predictive Analytics for Operations:** AI systems to forecast recruitment rates, identify potential bottlenecks, and dynamically reallocate resources or adjust trial sites. By analyzing trends, these tools could predict if a trial is likely to fail due to low accrual or detect when certain subgroups are lagging, enabling proactive mitigations.
- **Decentralized Trial Technologies:** AI-powered remote monitoring, telemedicine, and Internet-of-Things (IoT) devices to reduce patient burden and expand access. This includes algorithms to detect adherence or safety signals from wearable sensors or home-based measurements in real time. Embracing *decentralized clinical trials (DCTs)* – where visits can occur at local clinics or at home – requires robust AI to integrate data from these distributed sources and maintain data integrity. Federal investment can drive standards and validation for such AI so that regulators and sponsors confidently adopt DCT models.

Federal Role: Support the development of the underlying standards, infrastructure, and public goods needed to realize AI-accelerated trials. This includes establishing secure data sharing frameworks and federated learning networks that enable AI training on *diverse, real-world patient data* from multiple institutions (since no single trial sponsor has access to all patients). It also means convening stakeholders (academia, providers, patient advocacy groups, industry) to create open datasets and challenge problems for AI in trial design and enrollment. The government should prioritize R&D that ensures these AI tools are inclusive and generalizable – for instance, by requiring that any federated trial-matching AI be trained on data that reflects the nation’s demographic and genetic different populations. By catalyzing progress in AI for faster, fairer clinical trials, we can lower the cost of drug development, bring new treatments to patients sooner, and **broaden participation** so that scientific advances benefit *all* communities. The payoff will be a clinical research enterprise that is both more efficient and more equitable.

Robust AI for Real-World Evidence (RWE) Generation and Analysis

Need: Outside of trials, real-world clinical data – electronic health records (EHRs), claims, registries, patient-reported outcomes – holds immense potential to inform us about treatment effectiveness and safety in routine practice. Real-world evidence can complement trials by reflecting larger, more diverse populations and longer-term outcomes. However, mining reliable insights from these heterogeneous and often unstructured datasets is technically challenging. Issues of data quality, missing information, bias and confounding variables, and lack of standardization have hindered the use of RWE for high-stakes decisions. Sophisticated AI/ML techniques are needed to parse messy clinical text, control for myriad variables, and infer causal relationships (e.g. did a medication actually improve outcomes, or were those patients healthier to begin with?). Developing these methods is resource-intensive and not the primary focus of most companies, yet the public health benefit of trustworthy RWE is enormous – for example, the FDA can use RWE to approve new indications of drugs or to monitor post-market safety, and clinicians can learn how treatments work in populations not well represented in trials.

Priority: Invest in AI methodologies purpose-built for real-world biomedical data analysis. This

includes:

- **Causal Inference and Bias Mitigation:** Techniques like causal machine learning, propensity score modeling with AI, and bias detection algorithms that can handle the complexity of healthcare data (with its confounders and selection biases). Research should focus on AI that can *quantify uncertainty* and flag when correlations might be spurious. Federal funding can drive the creation of open benchmarks (e.g. simulated healthcare datasets with known causal ground truth) to evaluate and improve these algorithms.
- **Natural Language Processing (NLP) at Scale:** Much of real-world data is locked in free-text doctor's notes, pathology reports, and imaging narratives. Advancements in NLP – especially healthcare-tuned language models – are needed to extract clinically meaningful information (symptoms, adverse events, disease progression) reliably at scale. For instance, understanding that “patient’s tumor progressed on Drug X” in an oncologist’s note is crucial for outcome analysis. We recommend dedicated support for **clinical NLP** research, including multilingual capabilities, given the diversity of patient populations.
- **Data Harmonization and Integration:** AI tools that can automatically map and normalize data from different sources (different EHR vendors, pharmacy records, wearables) into a common representation. Building on standards like FHIR, the government can fund projects that create “AI-ready” datasets by cleaning and linking disparate data, so that researchers spend less time data-wrangling and more time on analysis.

By advancing AI for RWE, we enable continuous learning in healthcare. This is critical for regulatory science (e.g., evaluating drug performance outside controlled trials) and for providers making data-driven decisions. A robust federal push here will accelerate the development of learning health systems where data from every patient, in every setting, can inform best practices and innovation. In cancer care, for example, unlocking RWE with AI could reveal which treatments work best for which subgroups, or flag safety issues earlier. These are outcomes squarely in the public interest and deserving of federal R&D support.

Explainable, Trustworthy, and Fair AI in Healthcare

Need: The “black box” nature of many AI models is a major barrier to adoption in high-stakes clinical decision-making. Doctors, regulators, and patients are understandably cautious about algorithms that cannot explain their rationale. If an AI recommends a cancer treatment or flags a patient as high-risk, stakeholders need to understand **why** – both to trust the recommendation and to verify it aligns with clinical logic and ethical standards. Additionally, biases in AI systems (often reflecting biases in training data or societal disparities) can inadvertently perpetuate or even worsen healthcare inequities. There is mounting evidence of this risk: for example, a widely used health management algorithm was found to systematically underestimate the needs of Black patients, because it was trained on healthcare cost data that under-reflects the care Black patients receive ([Racial Bias Found in a Major Health Care Risk Algorithm | Scientific American](#)). Such biases can deny certain groups access to necessary interventions, exacerbating disparities. Ensuring AI is transparent, interpretable, and fair is not just an academic concern – it is essential for patient safety, ethical integrity, and public acceptance of AI in medicine. Yet, developing explainable and bias-resistant AI often requires fundamental advances (and interdisciplinary work with ethicists and social scientists) that profit-driven development may not prioritize.

Priority: Fund research and development of AI methods that are **inherently interpretable** or offer robust explanations for their outputs, as well as techniques to audit and ensure fairness. Key efforts should include:

- **Interpretable Models and XAI Tools:** Support the creation of algorithms that are more transparent by design (e.g., causal models, rule-based learners, generalized additive models with pairwise interactions) and **post-hoc** explanation techniques (like saliency maps, concept extraction, and natural language explanations) tailored for clinical use. These techniques must be evaluated in real clinical contexts – for instance, can an AI explain its prediction for risk of sepsis in a way that a bedside nurse and physician find useful and actionable? The federal government can establish benchmarks and testbeds to assess explainability methods on healthcare tasks, driving competition to improve them.
- **Fairness and Bias Evaluation Frameworks:** Develop standardized protocols and toolkits to detect bias in clinical AI systems – for example, evaluating model performance across race, gender, socioeconomic status, and other key demographics. Federal agencies (in collaboration with groups like the NIH and FDA) could require that any AI used in healthcare undergo such an evaluation. We recommend funding research into bias mitigation strategies, such as re-sampling techniques, fairness-constrained machine learning, or adversarial debiasing methods that can be applied during model training. Importantly, fairness is not just a technical metric but a societal one – thus support should also go to **community engagement and participatory design** approaches, bringing diverse patient perspectives into AI development.
- **Human-AI Collaboration:** Ensure that AI in healthcare is developed with a human-centered approach, aiming to *augment* rather than replace clinical judgment. This means researching how clinicians actually interact with AI recommendations, and how to present AI outputs in a user-friendly, confidence-calibrated manner. It may involve new interfaces (e.g., an AI “consultant” that can answer follow-up questions about a recommendation) or workflow integrations so that using the AI is seamless rather than burdensome. Cross-disciplinary research (bringing together AI designers, cognitive psychologists, and health professionals) is needed to get this right.

By making AI **explainable and worthy of trust**, we will unlock its full potential in medicine. Physicians are far more likely to use a decision support AI if it provides clear reasoning or supporting evidence, just as patients will trust AI-influenced care if they know it’s been vetted for fairness and accuracy. These are areas where federal R&D investment can set guardrails and quality standards, ensuring that as AI adoption grows, it does so in a way that upholds our values of transparency, equity, and accountability. In essence, we must strive for AI that doctors can comfortably “talk to” and validate, and that we as a society are confident will *do no harm* and indeed help **close gaps** in care rather than widen them.

AI for Public Health and Health Security

Need: Beyond individual patient care, AI can be a game-changer for public health surveillance, epidemiology, and the security of our healthcare infrastructure. The COVID-19 pandemic showed both the promise of data-driven outbreak modeling and the gaps in our preparedness. AI could help detect infectious disease outbreaks earlier, predict their spread, and optimize response (such as vaccine distribution) – literally saving lives and economic costs. It can also strengthen health

system resilience by optimizing supply chains for critical medical supplies or predicting hospital capacity needs during surges. However, these applications often fall outside the primary profit motives of industry; they are classic public goods where federal leadership is needed. Additionally, some public health data (e.g., epidemiological, environmental) may be fragmented across government agencies or not immediately accessible to tech companies. A concerted federal effort can integrate these data and pose public health challenges that spur AI innovation. Health security is also national security: threats like engineered pathogens or bioterrorism, as well as natural pandemics, require AI tools for rapid detection and response. Ensuring we have cutting-edge AI for these scenarios is squarely a federal responsibility.

Priority: Support AI R&D focused on public health outcomes and preparedness:

- **Predictive Epidemiology:** Develop AI models that utilize diverse data streams – clinical reports, pharmacy sales, social media, climate and vector data, genomics of pathogens, even wastewater signals – to identify emerging health threats and forecast their trajectory. Investments should be made in federated data networks and privacy-preserving analytics so that, for example, hospitals can contribute de-identified data to national outbreak models in real time. Challenge programs (perhaps through NIH or CDC) could task AI teams with improving early detection of specific diseases (flu, COVID variants, etc.) by X%.
- **Resource Allocation and Supply Chain AI:** COVID taught us that supply chain bottlenecks (for PPE, ventilators, drugs) can cripple response. AI optimization models can forecast demand and manage distribution logistics for critical supplies during public health emergencies. Federally funded research could simulate various disaster scenarios and use AI to devise robust strategies for allocation of limited resources (beds, medicines) to minimize loss of life.
- **Biosecurity and Pandemic Response:** Support AI that can aid in the rapid development of medical countermeasures. For instance, AI-driven protein folding and generative models (like AlphaFold or others) can accelerate vaccine or therapeutic design by predicting viral protein structures or suggesting drug candidates. Also, anomaly detection AI on health data that might flag a possible bioterror event or novel pathogen faster than traditional surveillance. While some of this overlaps with defense, a coordinated health-focused effort is needed.

By prioritizing AI for public health, the Strategic Plan ensures that advances in AI benefit not only individual healthcare consumers but also strengthen our collective health security. This focus will maintain U.S. leadership in using AI for societal resilience – a domain where we cannot afford to lag. Importantly, many innovations here (like better disease detection algorithms) will spin back into everyday healthcare, improving, for example, early diagnosis of illness in clinical practice. Federal investment in this space epitomizes high-impact AI for the public good, saving lives on a large scale and safeguarding our communities against 21st-century health threats.

Secure Data Infrastructure and Privacy-Preserving Technologies

Need: Progress in all the above areas is fundamentally limited by the challenge of **data** – accessing large-scale, diverse, high-quality health data while rigorously protecting patient privacy and rights. AI thrives on data, and to develop robust models (especially in a country as diverse as

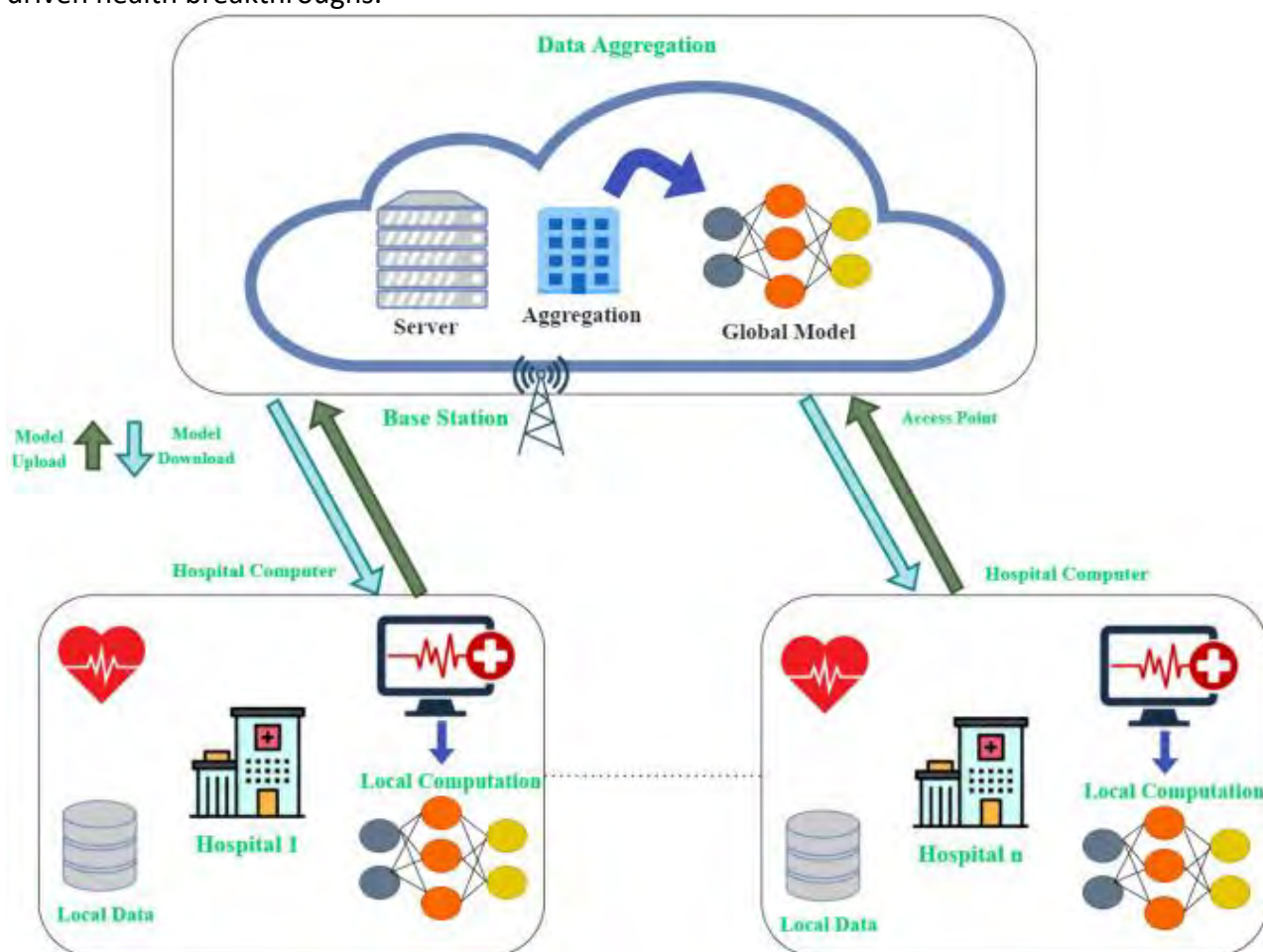
the U.S.), researchers need access to training data that reflect many populations and scenarios. Yet health data is rightly sensitive and heavily regulated (e.g. by HIPAA); it's siloed across thousands of providers and institutions; and even when accessible, often not in usable formats. Moreover, companies guard proprietary datasets as competitive assets, slowing collective progress. The federal government has a critical role to play in creating an environment where data can be shared or utilized **safely and ethically** at scale for AI R&D. This includes advancing technologies that enable data to be used without compromising privacy (so that individuals' medical information isn't exposed or misused) and establishing standards that make data from different sources "talk" to each other. In addition, generating **synthetic health data** – realistic but artificial patient records – offers a promising route to provide AI developers with abundant data without privacy issues. However, synthetic data generation methods need improvement to ensure realism and utility. Building a secure data infrastructure for health AI is akin to building roads and bridges for commerce: it is enabling infrastructure that the market alone will underprovide, but once in place, it benefits the entire ecosystem.

Priority: Invest in and coordinate the development of privacy-preserving data sharing technologies, synthetic data, and interoperability standards for health data. Key components:

- **Privacy-Enhancing Technologies (PETs):** These include *federated learning*, homomorphic encryption, secure multi-party computation, and differential privacy techniques, all tailored to health data scenarios. Federated learning allows multiple institutions to collaboratively train AI models on their local data, sending only model updates (gradients) to a central aggregator – so raw patient data never leaves the source site. Homomorphic encryption allows computations on encrypted data. Differential privacy adds noise to outputs to protect individual identities. The government should fund at-scale pilots and implementations of these PETs in healthcare, effectively creating a **national distributed data network** for research. A compelling example comes even from beyond Earth: NASA researchers recently demonstrated a federated learning framework between the International Space Station and Earth, training biomedical models without sending down raw data – instead exchanging secure model updates, even using synthetic data to supplement training. If such privacy-preserving AI is feasible in space, we can certainly deploy it across hospitals here on Earth. The Strategic Plan should make secure federated AI a cornerstone, so that major health systems, government agencies, and companies can *work together on AI models without ever pooling sensitive data*.
- **High-Fidelity Synthetic Health Data:** Support the development of tools to generate realistic synthetic patient datasets that can be used for AI training and testing. Synthetic data, if properly generated, can mimic the statistical properties of real clinical data without representing actual individuals. This can greatly lower barriers for AI development since it avoids privacy concerns. HHS's Synthea project is one example – it generates entire synthetic health records for hypothetical patients based on disease progression models. We recommend expanded efforts to create synthetic datasets for critical use cases (e.g., rare disease cohorts, multi-modal cancer patient journeys), and validation frameworks to evaluate how well AI trained on synthetic data performs on real data. Federal leadership can also clarify regulatory stances on using synthetic data (for instance, encouraging its use for algorithm benchmarking or software validation).
- **Interoperability and Data Standards:** Continue to drive standardization (building on HL7

FHIR, OMOP common data model, etc.) that makes disparate health data easier to aggregate and analyze. The goal should be a future where, with patient consent, data from any electronic health record system, lab, imaging device, or wearable can flow into research repositories in a consistent format. This may involve funding open-source tools for data mapping and cleaning, and enforcing compliance with interoperability rules for any federally funded health IT systems.

Investing in this data infrastructure yields a multiplier effect: it empowers all researchers, lowers the friction to develop and deploy AI, and ensures that privacy and security are baked into the foundation of health AI innovation. Without these enabling investments, efforts in other areas will be piecemeal and siloed. With them, we can create a **learning healthcare system** at national scale, where insights derived from millions of patients can rapidly inform care – all while keeping personal data safe. In summary, just as past generations invested in public utilities, we must now invest in the data infrastructure and privacy technology that will fuel American leadership in AI-driven health breakthroughs.



[Federated learning-based AI approaches in smart healthcare: concepts, taxonomies, challenges and open issues - PMC](#) Figure: Federated learning architecture for healthcare. In this approach, local nodes (e.g., Hospital 1...Hospital n) train AI models on **local data** and share only model parameters with a central aggregator in the cloud. The aggregator updates a global model and

sends improvements back to each node. This privacy-preserving design allows hospitals to collaboratively build powerful AI models without exchanging sensitive patient records. Such federated AI networks, combined with encryption and audit trails, are a cornerstone of the envisioned national infrastructure for health AI.

II. Novel Mechanisms to Accelerate AI R&D Through Partnerships

Achieving the above priorities will require not only funding specific research topics, but also **innovative partnership models** and policy mechanisms to unite stakeholders and translate research into practice. AI in healthcare sits at the intersection of multiple sectors – academia, government, healthcare providers, technology companies, startups, and patients. The federal government can play a catalytic role in bringing these parties together in pre-competitive collaboration and smoothing regulatory pathways. We propose several mechanisms to include in the strategic plan:

- **Federally-Facilitated Health Data Trusts & Collaboratives:** Establish pre-competitive consortia focused on pooling data and knowledge for AI research on specific diseases or public health challenges. For example, a *“National Oncology AI Data Collaborative”* could be formed to enable hospitals, research centers, patient advocacy groups, and industry partners (including companies like Massive Bio) to securely share data or to train federated models collectively. The government would provide governance, legal/compliance frameworks, and seed funding to ensure these collaboratives operate ethically and with privacy safeguards. By lowering traditional barriers (legal liability, competitive secrecy), such data trusts would unlock multi-institutional datasets of unprecedented scale. This would accelerate development of AI models that are more generalizable and powerful – e.g., an AI trained on oncology data from 100+ institutions rather than 5. The ARPA-H ACTR program we are part of is a microcosm of this approach, showing how a federated network can engage many partners toward a common goal. Making it a broader national endeavor – potentially through NIH or NSF – could bring similar models to other domains (cardiovascular, neurology, etc.). This allows progress without centralizing data, overcoming a major hurdle in health AI R&D.
- **“Grand Challenge” Initiatives for Healthcare AI:** Launch ambitious, milestone-driven prize competitions focused on specific high-impact AI applications. Modeled after past successes in driving innovation (such as DARPA’s Grand Challenges in autonomous vehicles), the government can define audacious but achievable targets – for instance, *“AI for Early Cancer Detection”* (develop an AI that can predict cancer at least 6 months before current methods for a certain type), or *“AI for Reducing Maternal Mortality Disparities”* (an AI system that demonstrably reduces the gap in maternal outcomes between different demographics). These challenges would come with funding, public recognition, and perhaps regulatory support for piloting the solutions. They signal national priorities and attract a wide pool of problem solvers, including non-traditional participants and interdisciplinary teams. By focusing energy on clear objectives, Grand Challenges can yield breakthrough prototypes and galvanize public excitement. Importantly, they also produce shared datasets and evaluation metrics (as teams need common test data to compete), which remain valuable community resources after the competition.

- **Regulatory Sandboxes for AI in Healthcare:** Create structured “safe spaces” where innovators can work closely with regulators (FDA, CMS, etc.) to test and validate AI technologies under real-world conditions and oversight. For example, a sandbox program could allow a hospital or company to deploy a new AI diagnostic tool in a limited clinical setting, with continuous monitoring and frequent feedback from FDA on what evidence would be needed for broader approval. This iterative approach would accelerate learning on both sides – developers better understand regulatory requirements and failure modes, and regulators gain familiarity with cutting-edge AI techniques, informing future guidelines. Such sandboxes would be time-bound and carefully monitored to protect patients, but would reduce the lengthy dead time in the current process where an AI might be developed in silico and not tested in clinical practice until very late. Sandboxes could also address **liability uncertainties** by, for instance, providing certain protections during the evaluation phase. We recommend OSTP and NITRD coordinate with FDA’s Digital Health Center of Excellence to implement at least one AI healthcare sandbox program within the next year or two. Areas like AI-driven medical imaging analysis or clinical decision support for sepsis could be good pilots. Ultimately, this will speed up the translation of AI research into safe and effective medical products, by providing a pragmatic bridge between the lab and widespread use.
- **AI Workforce Development for Healthcare:** Address the talent and knowledge gap by investing in the next-generation workforce that is fluent in both AI and health. This involves integrating AI training into medical and nursing education, as well as adding healthcare domain knowledge to data science curricula. For instance, federal grants could support the development of joint degree programs (MD/MS in Data Science, or PhD in AI with clinical rotations), internship programs bringing computer science students into hospitals and conversely clinicians into tech companies, and continuing education for current health professionals on AI tools. We also need more “translators” – people who understand both worlds and can facilitate adoption of AI in clinical settings. The plan should encourage interdisciplinary fellowships and career development awards at the intersection of AI, medicine, and public health. Furthermore, diversity in this workforce is key: ensuring opportunities for underrepresented communities in AI and biomedical data science will lead to more inclusive innovation. By building a workforce pipeline, we ensure the sustainability of our AI R&D initiatives and the ability to implement them broadly. In the near term, a national short-course curriculum on “AI for Healthcare” could be disseminated to all academic medical centers, supported by NIH/NSF, to rapidly raise the baseline expertise.

Each of these mechanisms – data collaboratives, grand challenges, sandboxes, and workforce programs – serves to knit together the strengths of government, academia, and industry in service of the public. They recognize that solving healthcare AI challenges is not just about algorithms, but about **ecosystems and people**. We urge the Strategic Plan to incorporate such innovative models, as they will greatly amplify the impact of federal R&D investments and shorten the path from research to real-world benefits.

III. Conclusion



Artificial intelligence holds unprecedented promise for transforming healthcare – from accelerating research, to improving clinical outcomes, to making care more efficient and accessible, and even bolstering our defense against public health crises. Realizing this promise will improve quality of life for millions of Americans, strengthen our economy, and ensure U.S. leadership in a critical technological frontier. Massive Bio is committed to being at the forefront of this transformation, particularly in oncology and clinical trials, where we have seen firsthand how AI can connect patients with lifesaving options that might otherwise be missed. But no single organization can unlock AI's full potential in healthcare alone. It requires strategic, sustained federal R&D investment focused on foundational challenges, enabling infrastructure, and those areas of market failure where only government coordination can align incentives with societal needs.

In this comment we have highlighted priority areas – from foundational AI research in biology, to AI for clinical trials, real-world evidence, explainability and fairness, public health applications, and the underpinning data infrastructure – where federal leadership is both necessary and highly impactful. We have also outlined mechanisms to accelerate progress through collaboration, sandboxes, and workforce training. Collectively, these recommendations chart a path for the 2025 National AI R&D Strategic Plan to **catalyze breakthroughs** that benefit all Americans. By prioritizing health and biomedicine within the national AI agenda, OSTP and NITRD will ensure that the coming era of AI innovation is one that directly translates into longer, healthier lives and a stronger society. This is a moment for bold vision: just as past federal initiatives led to the internet and GPS, a concerted initiative now can create an AI-enhanced healthcare system that seems miraculous by today's standards – curing diseases faster, tailoring treatments to individuals, and extending cutting-edge care to every corner of the country.

Massive Bio stands ready to collaborate with the government, academia, and other industry partners to advance this critical national priority. We are optimistic that, with the right investments and policies, the United States will remain the world leader in trustworthy AI for health, and in doing so, will secure a future where innovation and compassion go hand in hand.

We thank you for considering our input and for your commitment to harnessing AI for the public good.

Sincerely,

Dr. Arturo Loaiza-Bonilla, MD, MEd

Co-Founder & Chief Medical Officer

Massive Bio, Inc.

Chief of Hematology and Oncology, St Luke's University Health Network

Submitted on behalf of:

Massive Bio, Inc.

90 West Street #12M, New York, NY 10006, USA

Point of Contact:

- **Name:** [Dr. Arturo Loaiza-Bonilla](#), MD, MEd, FACP (Co-Founder & Chief Medical Officer)

Massive Bio, Inc., 2025

Approved for public dissemination and contains no confidential or business-proprietary information



We affirm that this document is **approved for public dissemination** and contains no confidential or business-proprietary information.