AI and ML for Precision Medicine

by

Navaneeth Malingan

AI and IoT Researcher, Developer, Educator

Nunnari(நுண்ணறி) Labs

NUNNARI LABS

AI Coimbatore

What I currently do?

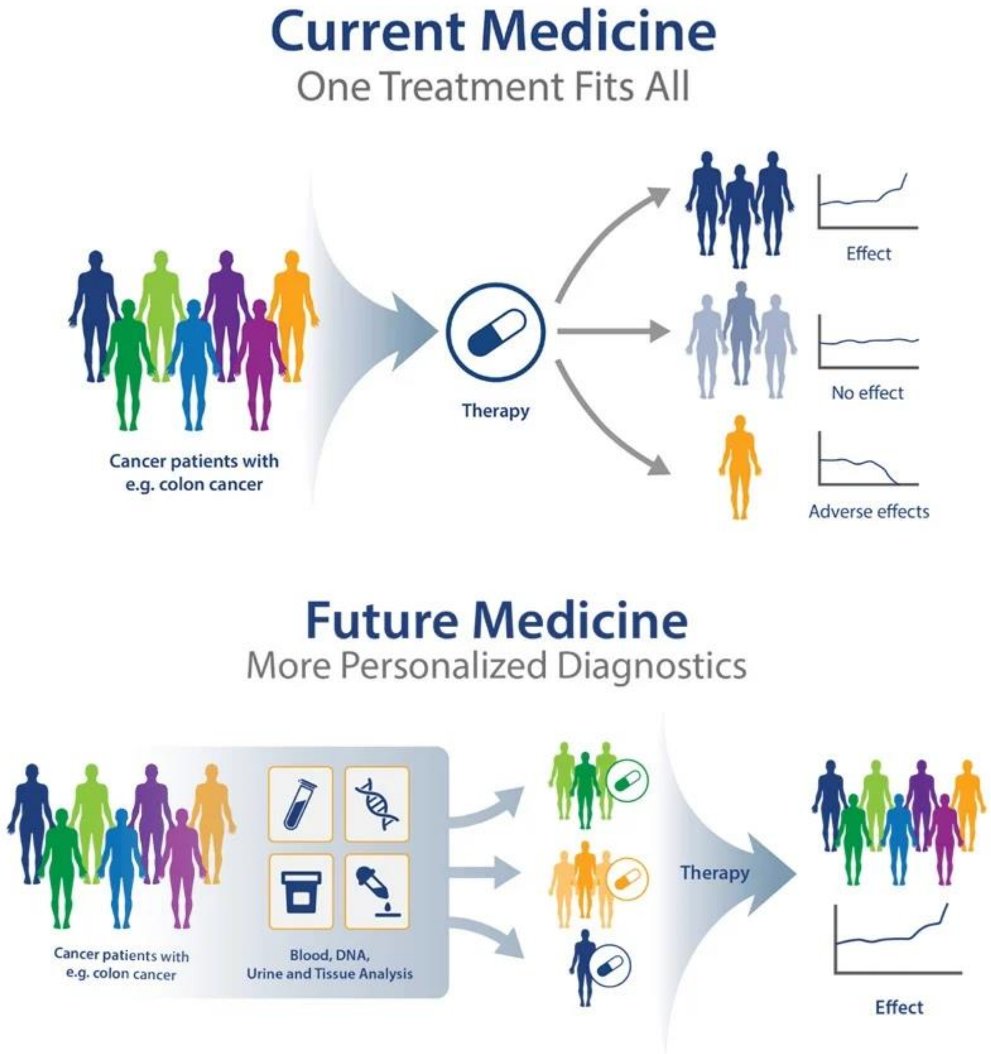
- Founder and AI Lead at Nunnari(நுண்ணறி) Labs
- Director – Innovations (KGiSL Educational Institutions)
- Technology Officer – Antpod (Agri), SeeDataMD (Healthcare)
- Community Lead – AI Tamil Nadu
- Adjunct Faculty - Ramakrishna Mission Vivekananda Educational and Research Institute and KCT

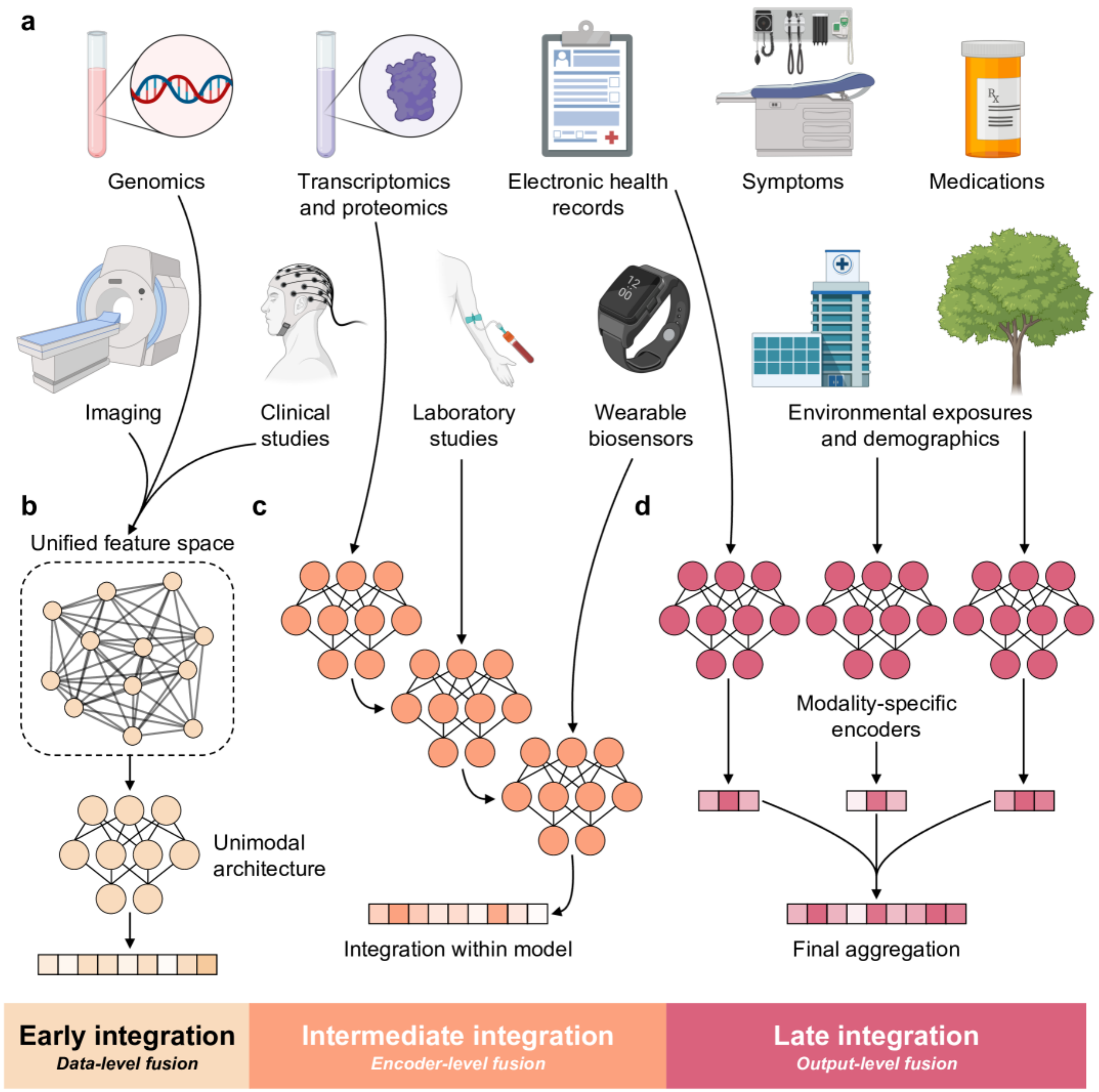
Slide 2: Introduction to Precision Medicine

- Definition and Overview
- Importance of Precision Medicine
- Historical Perspective and Evolution

Definition and Overview:

- **Precision Medicine**
  - A medical approach that customizes healthcare by **tailoring medical decisions**, **treatments**, **practices**, and **products** to **individual patients**.
  - Unlike the traditional **one-size-fits-all approach**, precision medicine considers individual variability in **genes**, **environment**, and **lifestyle**.





a
Genomics
Transcriptomics and proteomics
Electronic health records
Symptoms
Medications
Imaging
Clinical studies
Laboratory studies
Wearable biosensors
Environmental exposures and demographics
b
Unified feature space
Unimodal architecture
c
Integration within model
d
Modality-specific encoders
Final aggregation
Early integration
Data-level fusion
Intermediate integration
Encoder-level fusion
Late integration
Output-level fusion

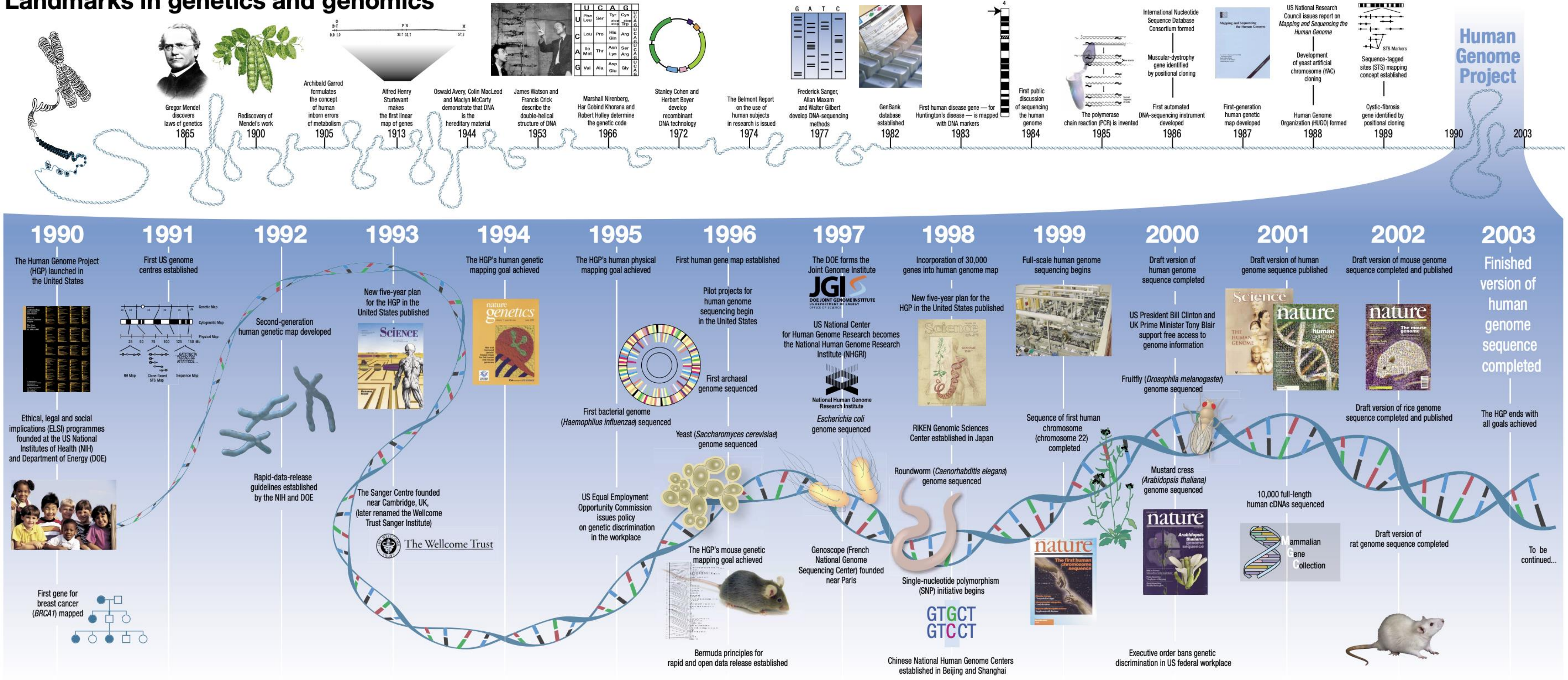
Importance of Precision Medicine:

- **Personalized Treatment:**
  - Treatments are designed based on the ***unique genetic*** makeup of each patient.
  - Increased effectiveness of treatments with ***fewer side effects***.
- **Enhanced Predictive Power:**
  - ***Better prediction*** of disease risk and response to treatment.
  - Early intervention and prevention strategies.
- **Improved Patient Outcomes:**
  - Higher ***success rates*** in treatments.
  - Increased ***patient satisfaction*** and quality of life.

Historical Perspective and Evolution:

- **Early Days:**
  - Medicine based on general observations and treatments that worked for the majority.
  - Discovery of DNA structure in 1953 laid the foundation for genetic research.
- **Milestones in Precision Medicine:**
  - **Human Genome Project (1990-2003):**
    - Mapping of the human genome revolutionized our understanding of genetics.
  - **Advances in Genomics and Biotechnology:**
    - Development of high-throughput sequencing technologies.
    - Rise of bioinformatics and computational biology.
  - **Introduction of AI and ML:**
    - Use of advanced algorithms to analyse vast amounts of genetic and clinical data.
- **Current State and Future Prospects:**
  - Ongoing research and clinical trials.
  - Integration of multi-omics data (genomics, proteomics, metabolomics).
  - Future potential to develop even more refined and effective personalized treatments.

Landmarks in genetics and genomics



PEAS COURTESY J. BLAMIRE, CITY UNIV. NEW YORK; WATSON & CRICK COURTESY A. BARRINGTON BROWN/SPL; *SCIENCE* COVERS COURTESY AAAS

Few Interesting Information from Human Genome Project

- There are about 20,000 to 25,000 genes in the human genome
- Human genes are distributed across 23 pairs of chromosomes (46 chromosomes in total), with each chromosome containing hundreds to thousands of genes.
- **BRCA1 and BRCA2:** These genes are involved in DNA repair and, when mutated, increase the risk of breast and ovarian cancer.
- **TP53:** This gene produces a protein that regulates the cell cycle and prevents cancer; mutations in TP53 are found in many types of cancer.

Slide 3: Role of AI and ML in Precision Medicine

- Why AI and ML are Crucial
- Key Benefits
- Overview of Impact Areas

Why AI and ML are Crucial:

- **Handling Vast Amounts of Data:**
  - Ability to process and analyze large-scale medical datasets efficiently.
  - Extracting meaningful insights from complex and heterogeneous data sources.
- **Identifying Patterns and Insights:**
  - AI and ML algorithms can detect patterns and correlations that are not evident to human researchers.
  - Enhancing the accuracy and speed of diagnosis and treatment recommendations.

Key Benefits:

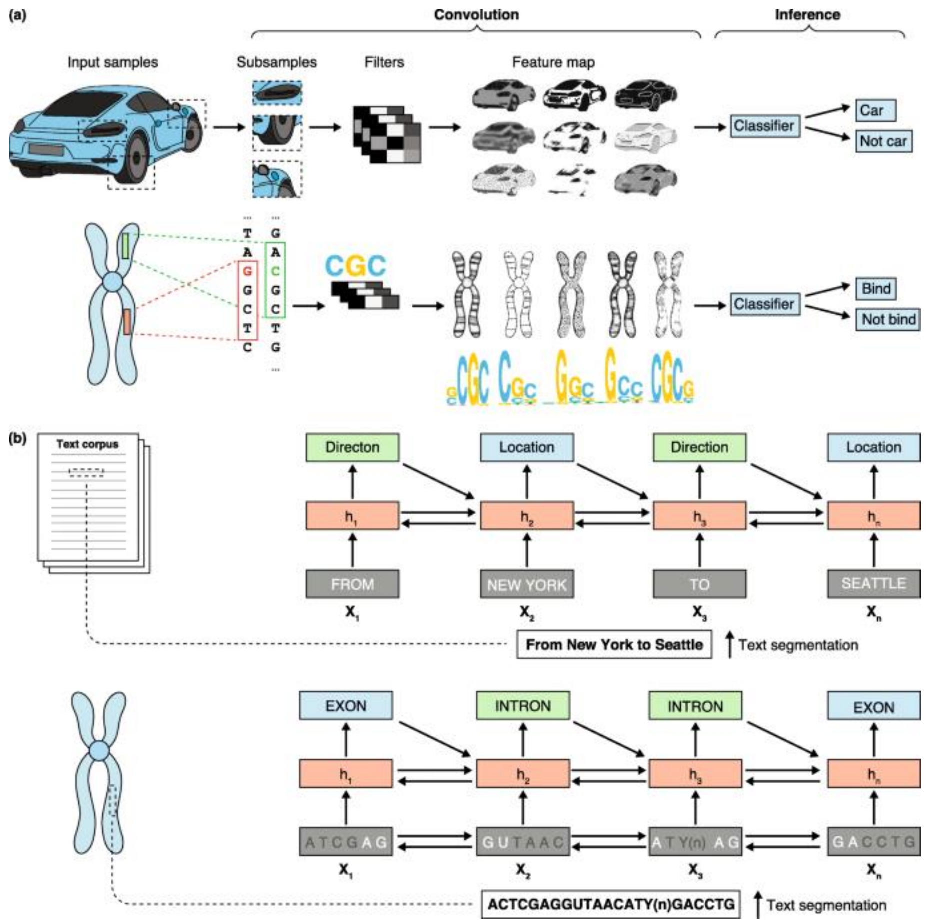
- **Improved Diagnostics Accuracy:**
  - Enhanced image analysis for radiology and pathology.
  - Early detection of diseases through predictive modelling.
- **Personalized Treatment Recommendations:**
  - Tailoring treatment plans based on individual genetic profiles and clinical histories.
  - Adaptive treatment strategies that evolve with patient response.
- **Predictive Analytics for Disease Progression:**
  - Forecasting disease trajectories and outcomes.
  - Proactive intervention strategies to prevent complications.

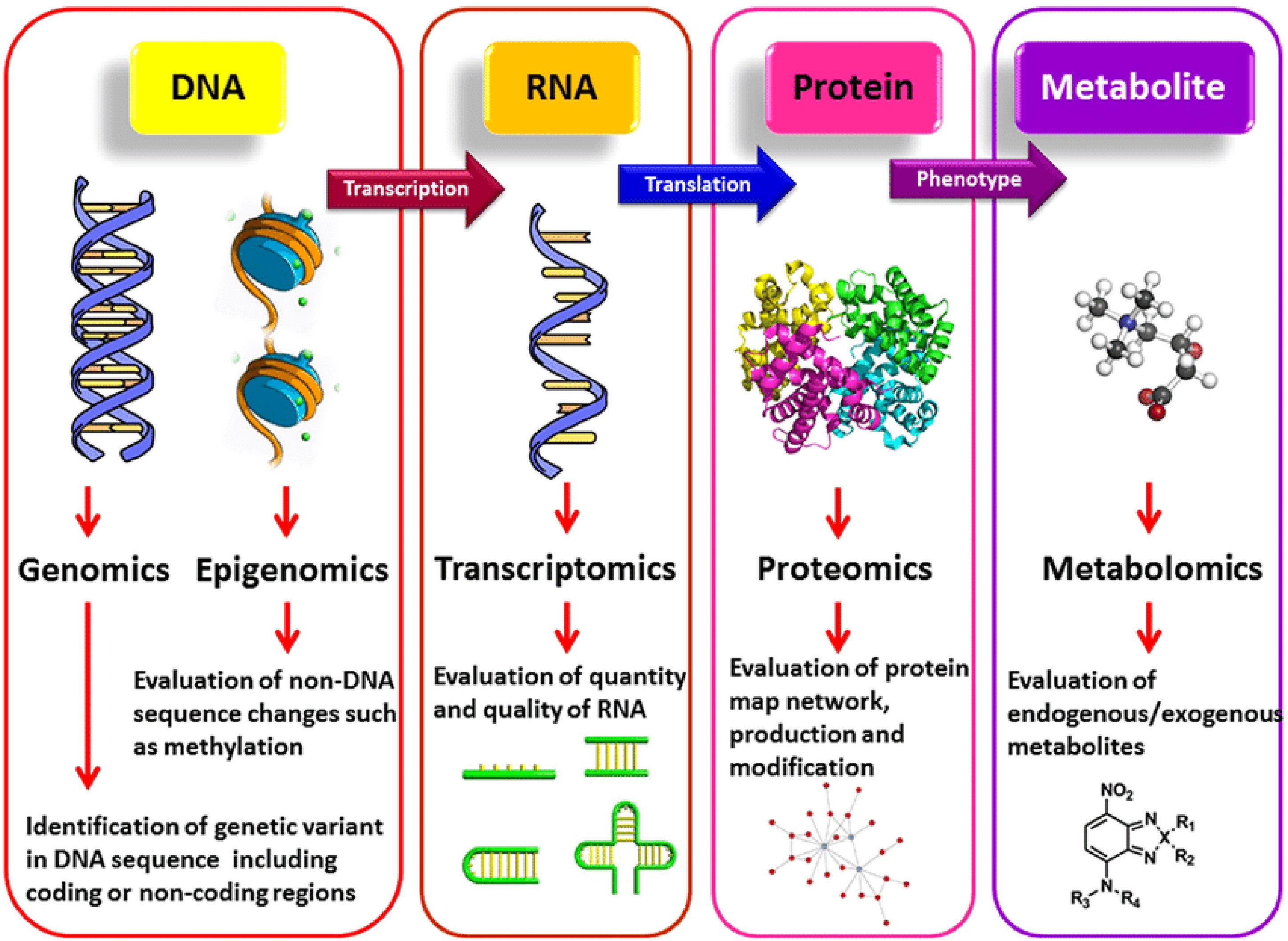
Overview of Impact Areas:

- **Genomics**
- **Imaging**
- **Pathology**
- **Treatment Plans**
- **Patient Monitoring**

Genomics

- AI-driven analysis of genomic data to identify genetic mutations and their implications.
- Personalized medicine based on genetic predispositions.




DNA
RNA
Protein
Metabolite
Transcription
Translation
Phenotype
Genomics
Epigenomics
Transcriptomics
Proteomics
Metabolomics
Evaluation of non-DNA sequence changes such as methylation
Evaluation of quantity and quality of RNA
Evaluation of protein map network, production and modification
Evaluation of endogenous/exogenous metabolites
Identification of genetic variant in DNA sequence including coding or non-coding regions

The Scale of Data in a Single Biopsy

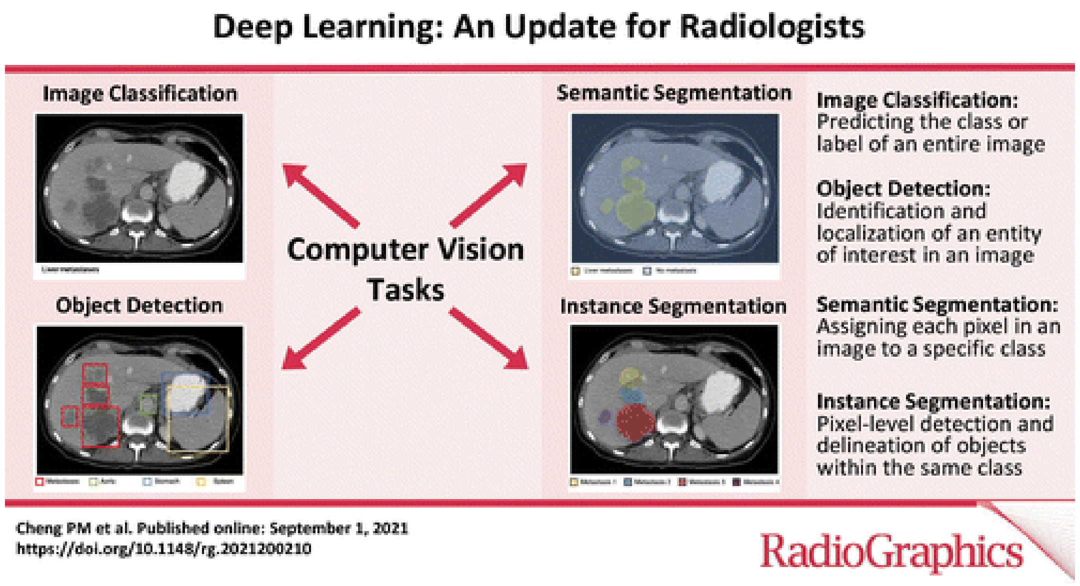
(2000 cells, 500 cancer cells)

Biopsy	Mutations	Genes	Proteins	Cells
1	100K	10M	200K	34M

Over <u>44 Million data points</u> per biopsy

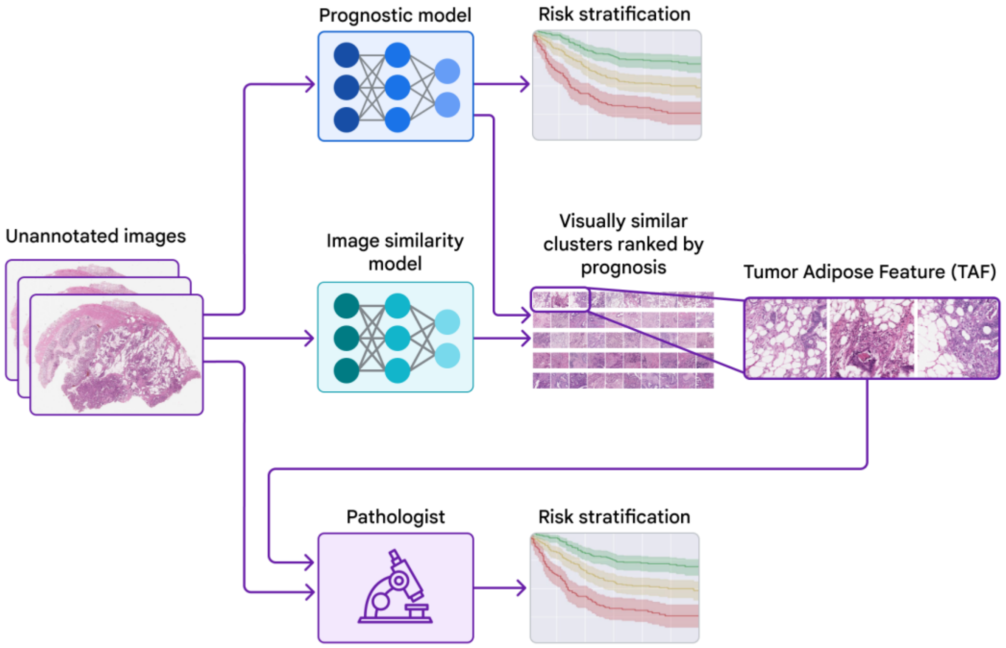
Imaging

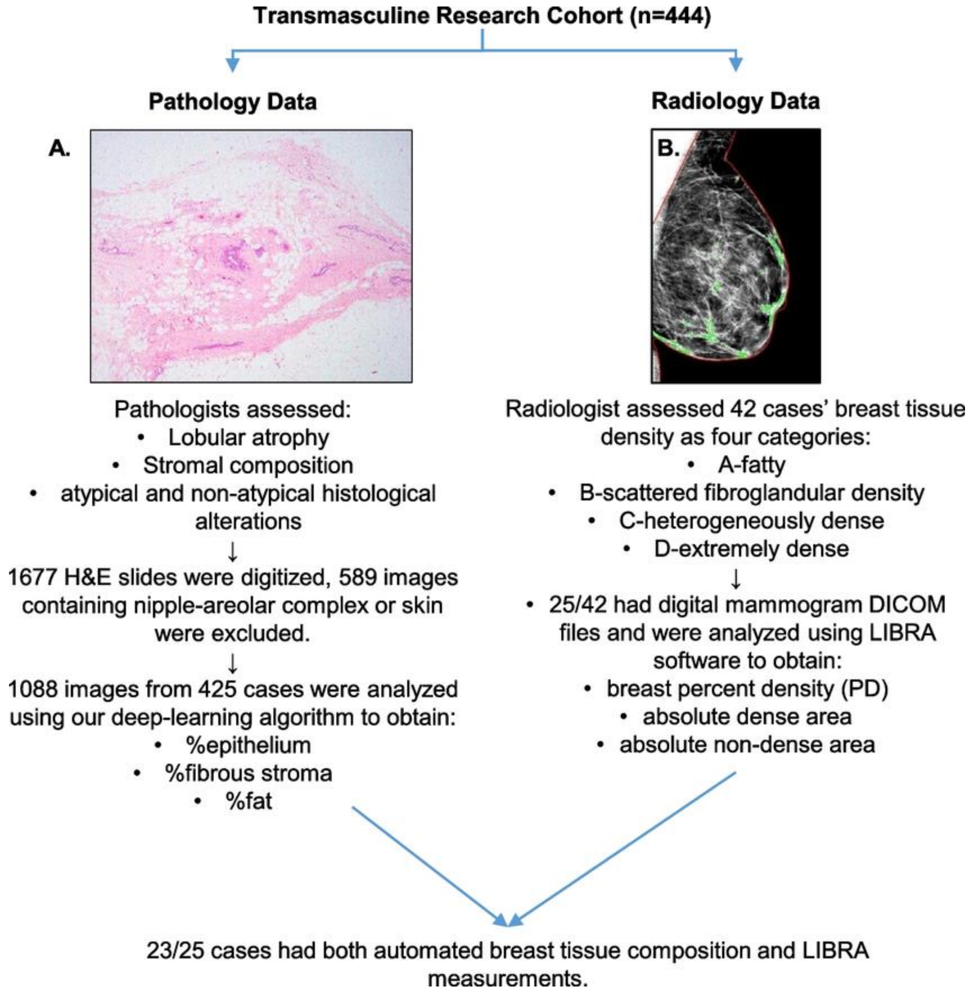
- Automated analysis of medical images (MRI, CT scans, X-rays) for accurate and early diagnosis.
- Reducing the workload for radiologists and increasing diagnostic precision.

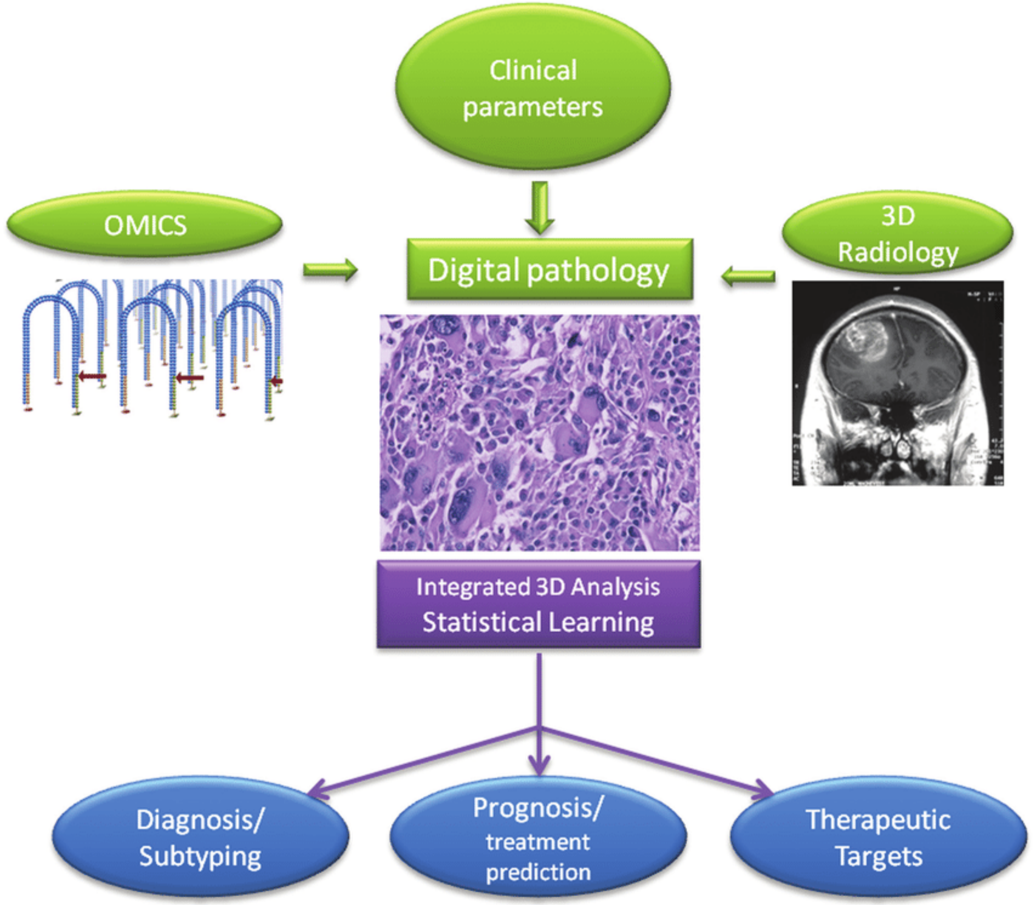


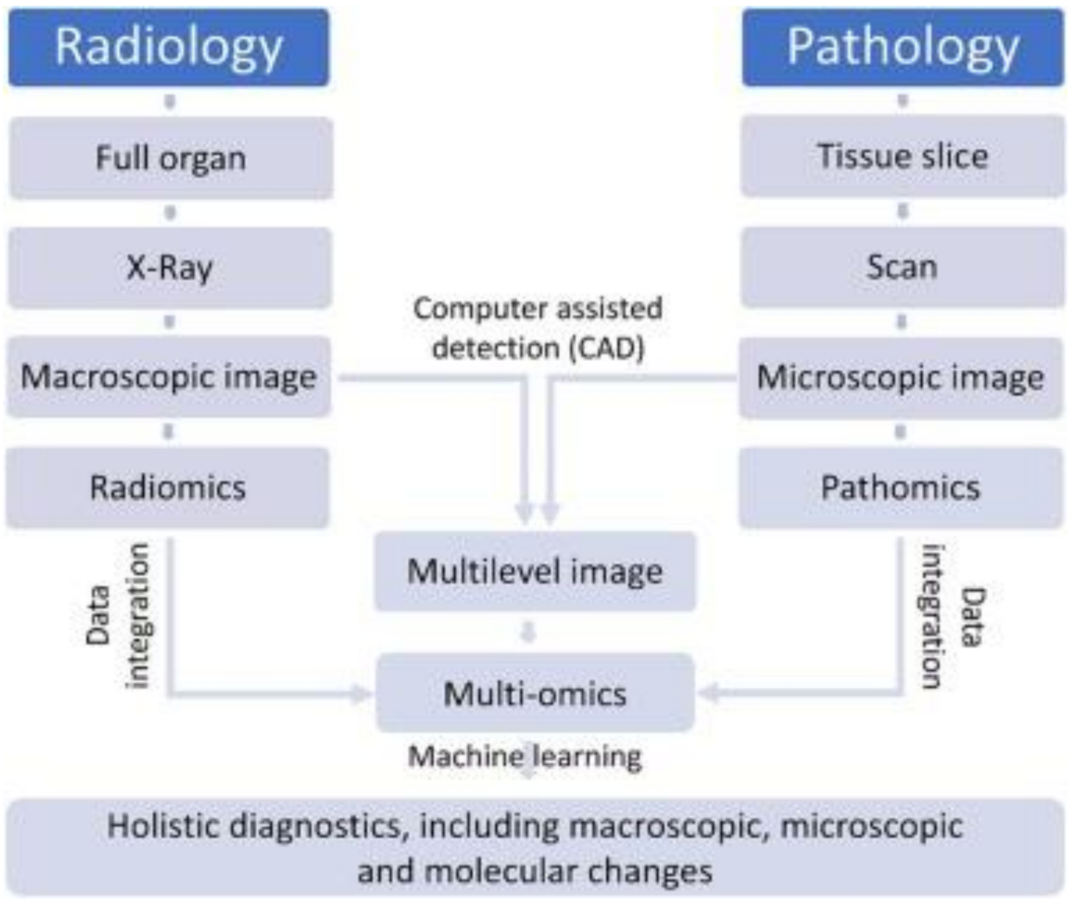
Pathology

- AI algorithms to analyse histopathological images for cancer detection and other diseases.
- Standardizing interpretations and reducing inter-observer variability.

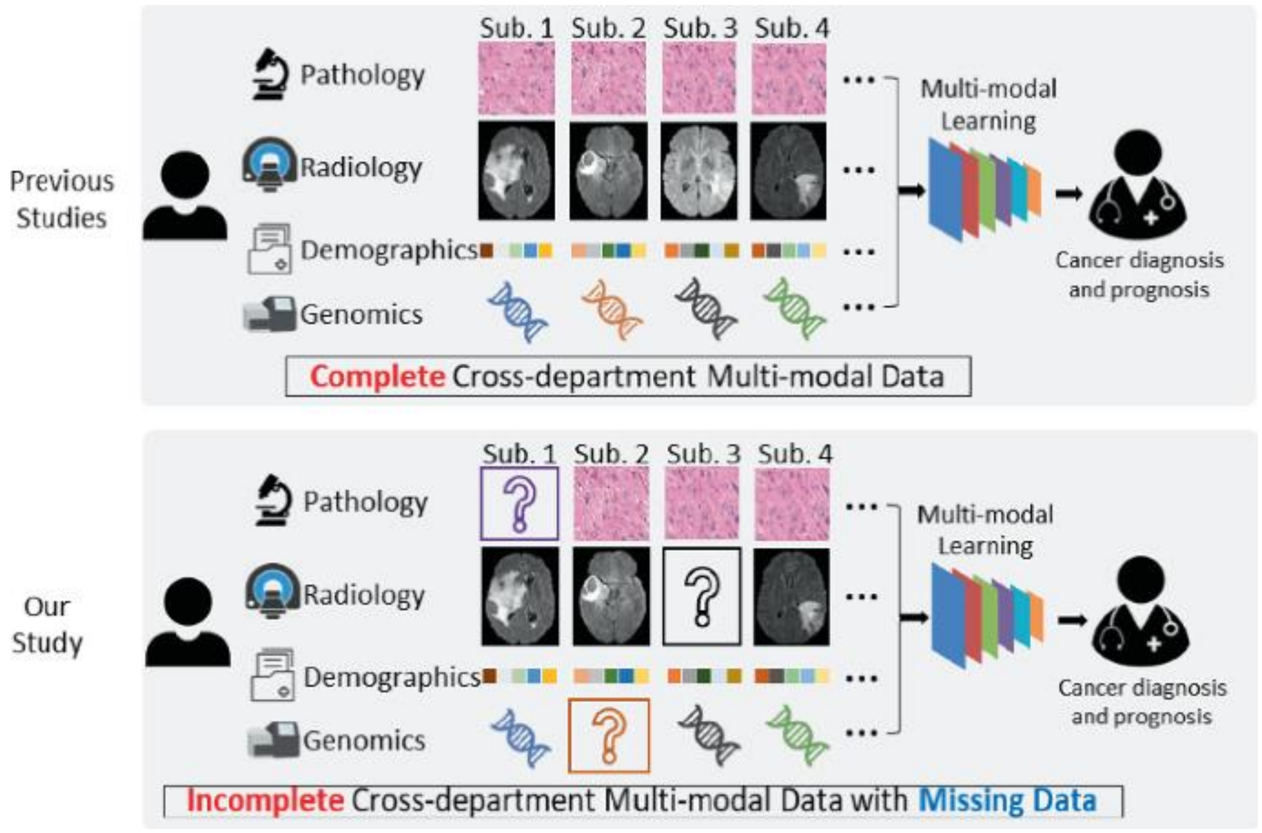



Transmasculine Research Cohort (n=444)
Pathology Data
Radiology Data
A.
B.
Pathologists assessed:
• Lobular atrophy
• Stromal composition
• atypical and non-atypical histological alterations
↓
1677 H&E slides were digitized, 589 images containing nipple-areolar complex or skin were excluded.
↓
1088 images from 425 cases were analyzed using our deep-learning algorithm to obtain:
• %epithelium
• %fibrous stroma
• %fat
Radiologist assessed 42 cases' breast tissue density as four categories:
• A-fatty
• B-scattered fibroglandular density
• C-heterogeneously dense
• D-extremely dense
↓
• 25/42 had digital mammogram DICOM files and were analyzed using LIBRA software to obtain:
• breast percent density (PD)
• absolute dense area
• absolute non-dense area
23/25 cases had both automated breast tissue composition and LIBRA measurements.


Clinical parameters
OMICS
Digital pathology
3D Radiology
Integrated 3D Analysis
Statistical Learning
Diagnosis/ Subtyping
Prognosis/ treatment prediction
Therapeutic Targets

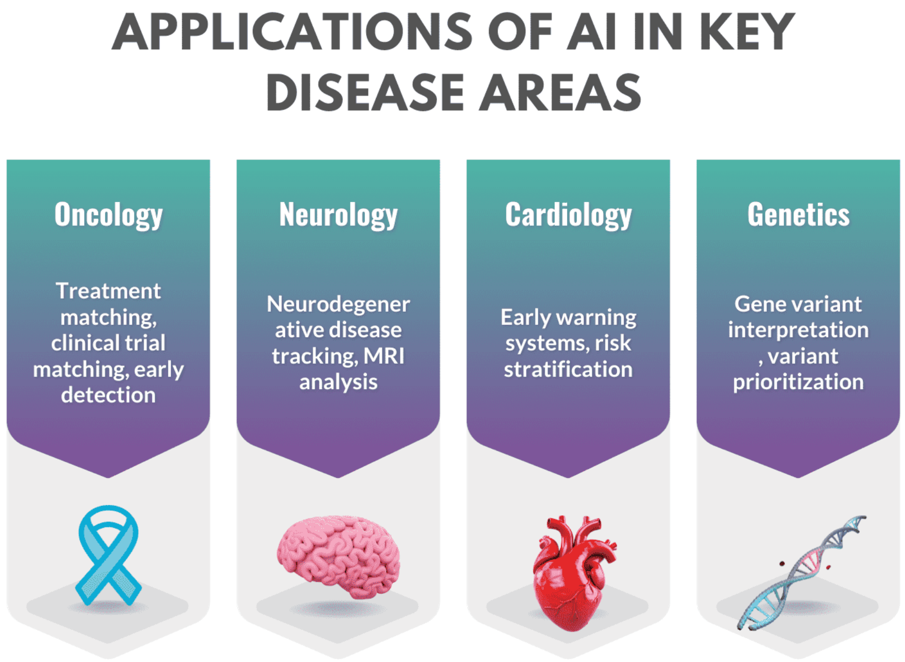

Radiology
Full organ
X-Ray
Macroscopic image
Radiomics
Pathology
Tissue slice
Scan
Microscopic image
Pathomics
Computer assisted detection (CAD)
Multilevel image
Data integration
Data integration
Multi-omics
Machine learning
Holistic diagnostics, including macroscopic, microscopic and molecular changes

Survival Prediction of Brain Cancer with Incomplete Radiology, Pathology, Genomic, and Demographic Data



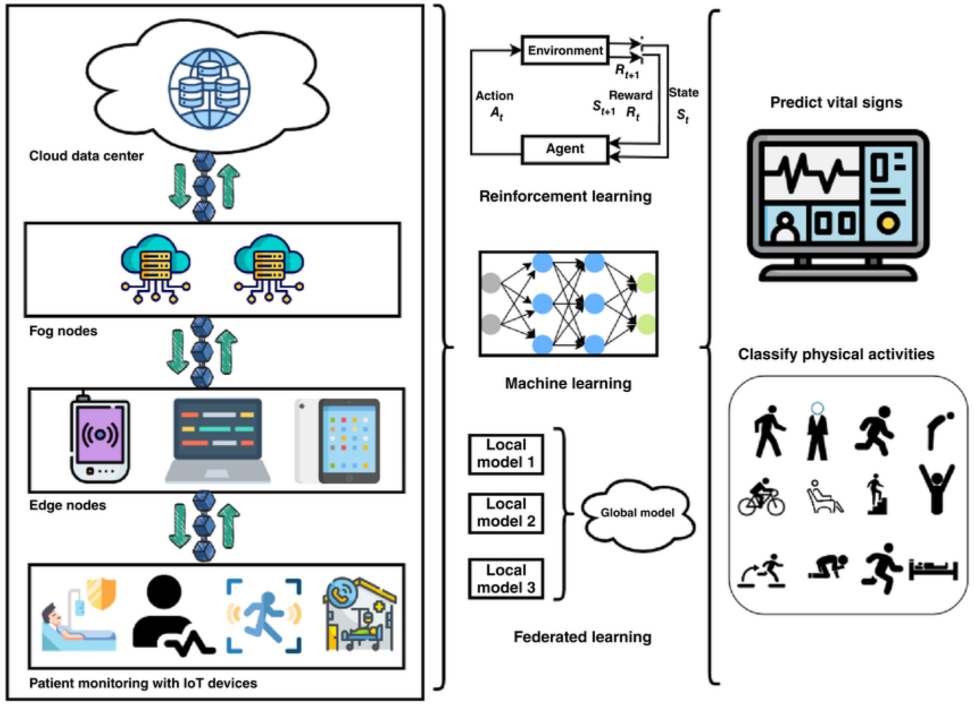
Treatment Plans

- Developing and optimizing personalized treatment regimens.
- Monitoring patient responses and adjusting treatments in real-time.



Patient Monitoring

- Continuous monitoring through wearables and IoT devices.
- Real-time data analysis to detect anomalies and provide timely interventions.



Slide 4: Data Acquisition and Management

- Sources of Medical Data
  - Genomics
  - Electronic Health Records (EHR)
  - Wearables and IoT Devices
- Data Privacy and Security
- Data Preprocessing Techniques

Sources of Medical Data

- **Genomics:**
  - **DNA Sequencing:**
    - Techniques like Next-Generation Sequencing (NGS) provide comprehensive genetic data.
    - Whole-genome sequencing, exome sequencing, and targeted gene panels.
  - **Genetic Markers:**
    - Single Nucleotide Polymorphisms (SNPs), copy number variations (CNVs), and other genetic variations.
- **Electronic Health Records (EHR):**
  - **Patient Histories:**
    - Longitudinal records of patient health over time.
  - **Clinical Data:**
    - Data from hospital visits, lab tests, diagnoses, medications, and treatment plans.
  - **Structured and Unstructured Data:**
    - Structured: Coded entries such as ICD codes.
    - Unstructured: Clinical notes, free-text entries.
- **Wearables and IoT Devices:**
  - **Continuous Monitoring:**
    - Devices like smartwatches, fitness trackers, and medical-grade sensors.
  - **Data Types:**
    - Heart rate, activity levels, glucose levels, sleep patterns, etc.

Data Privacy and Security

- **Importance of Maintaining Patient Confidentiality:**
  - Legal and ethical obligation to protect patient data.
  - Building trust with patients and healthcare providers.
- **Security Measures:**
  - **Encryption:**
    - Encrypting data at rest and in transit to prevent unauthorized access.
  - **Access Control:**
    - Role-based access controls to limit data access to authorized personnel.
  - **Anonymization and De-identification:**
    - Removing personally identifiable information (PII) to protect patient privacy.
- **Compliance with Regulations:**
  - **HIPAA (Health Insurance Portability and Accountability Act):**
    - U.S. regulations protecting patient health information.
  - **GDPR (General Data Protection Regulation):**
    - European Union regulations on data protection and privacy.

Data Preprocessing Techniques

- **Cleaning and Normalization:**
  - **Data Cleaning:**
    - Identifying and correcting errors, inconsistencies, and missing values.
  - **Normalization:**
    - Scaling data to a standard range to improve algorithm performance.
- **Handling Missing Data:**
  - **Imputation Techniques:**
    - Filling missing values using statistical methods, machine learning models, or domain knowledge.
- **Handling Outliers:**
  - **Outlier Detection:**
    - Identifying and treating outliers that can skew analysis and model training.
- **Feature Engineering:**
  - **Creating New Features:**
    - Deriving meaningful features from raw data to improve model performance.
  - **Feature Selection:**
    - Selecting the most relevant features to reduce dimensionality and improve efficiency.

Slide 5: Data Privacy, Security, and Ethical Considerations

- Importance of Data Privacy
- Security Measures
- Ethical Considerations

Importance of Data Privacy

- **Protecting Patient Information:**
  - Ensuring that sensitive patient information is not exposed or misused.
  - Maintaining patient trust and the integrity of the healthcare system.
- **Legal and Ethical Obligations:**
  - Healthcare providers and researchers are legally required to protect patient data.
  - Ethical responsibility to respect patient autonomy and confidentiality.

Security Measures

- **Encryption:**
  - **Data Encryption at Rest:**
    - Encrypting stored data to prevent unauthorized access in case of data breaches.
  - **Data Encryption in Transit:**
    - Securing data as it moves between systems, preventing interception during transmission.
- **Access Control:**
  - **Role-based Access Control (RBAC):**
    - Granting access to data based on user roles and responsibilities.
  - **Multi-Factor Authentication (MFA):**
    - Adding an extra layer of security by requiring multiple forms of verification.
- **Anonymization and De-identification:**
  - **Removing Personally Identifiable Information (PII):**
    - Techniques to strip data of information that can directly identify individuals.
  - **Pseudonymization:**
    - Replacing private identifiers with fake identifiers or pseudonyms.
- **Audit Trails and Monitoring:**
  - **Logging Access and Changes:**
    - Keeping detailed logs of who accessed data, when, and what changes were made.
  - **Real-time Monitoring:**
    - Implementing systems to monitor data access and flag suspicious activities

Compliance with Regulations

- **HIPAA (Health Insurance Portability and Accountability Act):**
  - **Overview:**
    - U.S. regulation that provides data privacy and security provisions to safeguard medical information.
  - **Key Requirements:**
    - Standards for electronic health transactions, privacy rule, security rule.
- **GDPR (General Data Protection Regulation):**
  - **Overview:**
    - Regulation in the European Union on data protection and privacy.
  - **Key Requirements:**
    - Right to access, right to be forgotten, data portability, breach notification.
- **Other Regional Regulations:**
  - **CCPA (California Consumer Privacy Act):**
    - U.S. state-level regulation providing privacy rights to California residents.
  - **PIPEDA (Personal Information Protection and Electronic Documents Act):**
    - Canadian law governing data privacy.

Ethical Considerations

- **Consent and Transparency:**
  - **Informed Consent:**
    - Ensuring patients understand how their data will be used and have consented to it.
  - **Transparency:**
    - Clear communication about data usage, storage, and sharing practices.
- **Addressing Biases in AI Models:**
  - **Bias Identification:**
    - Recognizing and addressing biases that may exist in training data and algorithms.
  - **Fairness and Inclusivity:**
    - Ensuring AI models are fair and do not disproportionately affect certain groups.
- **Data Ownership and Control:**
  - **Patient Rights:**
    - Patients' rights to control their data and decide how it is used.
  - **Data Portability:**
    - Patients' ability to transfer their data between providers and services.

Slide 6: AI and ML Techniques in Precision Medicine

- Overview of Common Algorithms
    - Supervised Learning
    - Unsupervised Learning
    - Deep Learning
- Applications in Diagnostics, Prognostics, and Treatment
- Case Studies

Overview of Common Algorithms

- **Supervised Learning:**
  - **Decision Trees and Random Forests:**
    - Used for classification and regression tasks, such as ***predicting disease outcomes*** based on patient data.
  - **Support Vector Machines (SVM):**
    - Effective for ***high-dimensional spaces***, commonly used in image classification.
  - **Neural Networks:**
    - **Deep Learning (DL):**
      - Convolutional Neural Networks (CNNs): Primarily used for image analysis in ***radiology*** and ***pathology***.
      - Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM): Used for time-series data analysis, such as ***patient monitoring and EHR analysis***.
  - **K-Nearest Neighbours (KNN):**
    - Simple algorithm for classification tasks, such as ***identifying similar patient cases***.
  - **Linear and Logistic Regression:**
    - Fundamental algorithms for ***risk prediction and binary classification tasks***.

Unsupervised Learning:

- **Clustering Algorithms (K-Means, Hierarchical Clustering):**
  - Grouping patients with similar characteristics for targeted treatments.
- **Principal Component Analysis (PCA):**
  - Dimensionality reduction for genomic data analysis.
- **Autoencoders:**
  - Used for anomaly detection and data compression.

Reinforcement Learning:

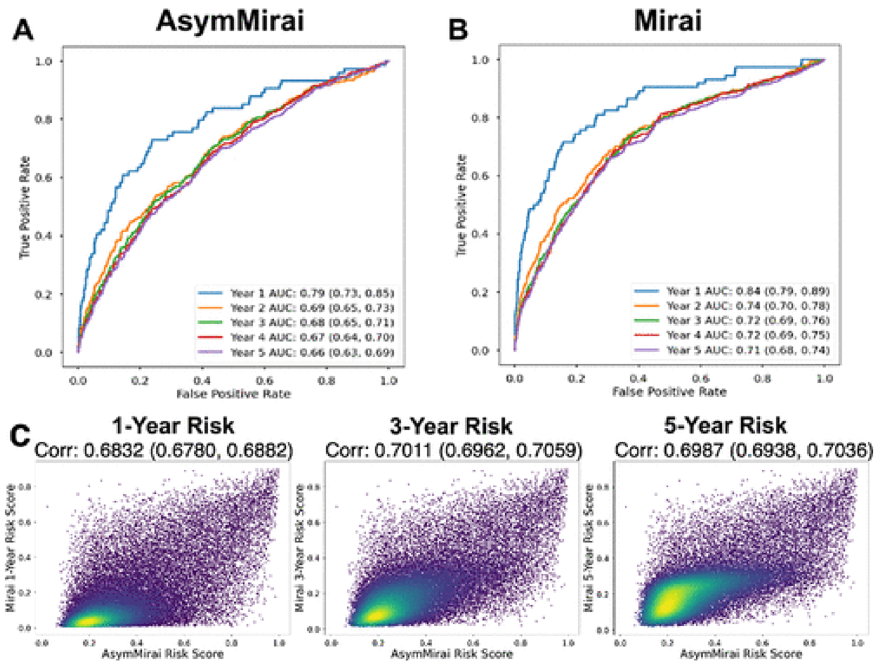
- **Q-Learning and Deep Q-Networks (DQNs):**
  - Developing adaptive treatment strategies that learn optimal policies over time.

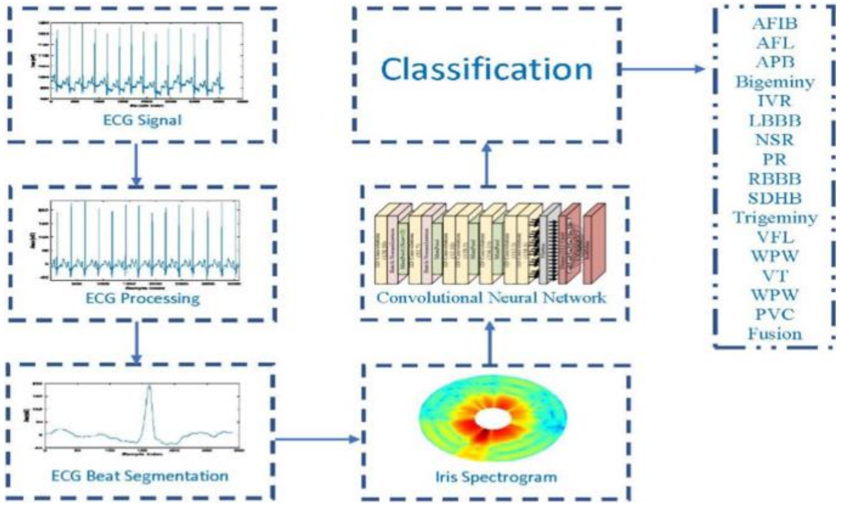
Applications in Diagnostics, Prognostics, and Treatment

- **Diagnostics:**
  - Image analysis for detecting tumours and other anomalies (e.g., CNNs for radiology).
  - Predictive models for identifying high-risk patients (e.g., Random Forests for disease prediction).
- **Prognostics:**
  - Predicting disease progression and patient outcomes (e.g., LSTM networks for time-series data).
  - Survival analysis using Cox Proportional Hazards Model.
- **Treatment Recommendations:**
  - Personalized medicine based on genetic profiles (e.g., Decision Trees for treatment pathways).
  - Adaptive treatment strategies using Reinforcement Learning.

Case Studies:

- **Cancer Detection:**
  - Using CNNs for early detection of breast cancer from mammograms.
  - AI models predicting cancer recurrence based on patient history and genetic markers.
- **Cardiology:**
  - SVMs and Neural Networks for identifying cardiac arrhythmias from ECG data.
  - Predicting heart disease risk using logistic regression on patient lifestyle and medical history data.





Slide 7: Case Studies in Diagnostics

- Example 1: AI in Cancer Detection
- Example 2: Predicting Heart Disease

Disease	Objective	Dataset Type	Types of AI Algorithms	Impact
Breast Cancer	Early detection of breast cancer through mammograms	Mammogram images	Convolutional Neural Networks (CNNs)	Improved early detection rates; reduced false positives and negatives
Lung Cancer	Detect lung nodules and classify them as benign/malignant	CT scan images	Convolutional Neural Networks (CNNs)	Accurate identification of malignant nodules; timely intervention
Heart Disease	Predict the risk of heart disease	Patient demographics, lifestyle, clinical history, genetic data	Logistic Regression	Preventive measures for at-risk individuals; reduced incidence of heart disease
Cardiac Arrhythmias	Detect cardiac arrhythmias from ECG data	ECG recordings	Support Vector Machines (SVMs), Neural Networks	Improved diagnosis and management of arrhythmias; enhanced patient monitoring

Results

- **Breast Cancer Detection:**
  - Example mammogram images with highlighted regions of interest (ROIs) detected by AI.
  - Graphs showing performance metrics (accuracy, sensitivity, specificity) before and after AI implementation.
- **Lung Cancer Detection:**
  - CT scan images with AI-detected lung nodules.
  - Performance metrics and comparison with traditional diagnostic methods.
- **Heart Disease Prediction:**
  - ROC curve illustrating the model's predictive accuracy.
  - Infographics on patient data used for prediction.
- **Cardiac Arrhythmias Detection:**
  - ECG graphs with annotated arrhythmia types identified by AI.
  - Confusion matrix showing model performance.

Slide 8: Personalized Treatment Plans

- Tailoring Treatment Plans
- Predictive Analytics for Disease Progression
- Challenges and Limitations

Tailoring Treatment Plans

- **Overview:**
  - Personalized treatment plans are designed to address the unique needs of individual patients based on their genetic, clinical, and lifestyle data.
  - Leveraging AI and ML to analyse complex datasets and create customized treatment protocols.
- **Steps in Creating Personalized Treatment Plans:**
  - **Data Collection:**
    - Comprehensive gathering of patient data, including genetic information, medical history, lifestyle factors, and current health status.
  - **Data Analysis:**
    - Using AI/ML algorithms to analyse collected data and identify patterns, correlations, and insights.
  - **Treatment Design:**
    - Developing a tailored treatment plan that considers the patient's unique characteristics and predicted response to various treatments.
  - **Implementation:**
    - Applying the personalized treatment plan in clinical practice, with continuous monitoring and adjustments as needed.

Predictive Analytics for Disease Progression

- **Importance of Predictive Analytics:**
  - Predictive models help forecast disease progression and patient outcomes, enabling proactive and preventive healthcare.
  - Early identification of potential complications allows for timely interventions and better management of chronic conditions.
- **Techniques Used:**
  - **Time-Series Analysis:**
    - LSTM networks and other time-series models to analyze sequential health data and predict future health events.
  - **Survival Analysis:**
    - Cox Proportional Hazards Model and other survival analysis techniques to estimate the time to an event, such as disease recurrence.
- **Applications:**
  - **Chronic Disease Management:**
    - Predicting disease exacerbations in conditions like diabetes, COPD, and heart failure.
  - **Cancer Treatment:**
    - Forecasting tumour growth and metastasis, guiding treatment adjustments.

Challenges and Limitations

- **Data Quality and Integration:**
  - Ensuring high-quality, comprehensive data from multiple sources.
  - Integrating diverse datasets (genomic, clinical, lifestyle) into a cohesive analysis framework.
- **Model Accuracy and Bias:**
  - Ensuring AI/ML models are accurate and generalizable.
  - Addressing potential biases in data and algorithms that could impact treatment recommendations.
- **Clinical Implementation:**
  - Integrating AI-driven personalized treatment plans into existing clinical workflows.
  - Training healthcare providers to effectively use AI tools and interpret their outputs.
- **Ethical and Regulatory Considerations:**
  - Navigating ethical issues related to patient privacy, data security, and informed consent.
  - Complying with regulatory requirements and standards for AI in healthcare.

Slide 9: Integration with Clinical Practice

- Integrating AI/ML Tools into Clinical Workflows
- Collaboration between Data Scientists and Healthcare Professionals
- Training and Support for Clinicians

Integrating AI/ML Tools into Clinical Workflows

- **Seamless Integration:**
  - Importance of incorporating AI/ML tools without disrupting existing clinical workflows.
  - Strategies for smooth integration include phased rollouts and comprehensive training programs.
- **Clinical Decision Support Systems (CDSS):**
  - AI-powered systems providing real-time support to clinicians.
  - Examples include diagnostic tools, treatment recommendation systems, and predictive analytics.
- **Electronic Health Records (EHR) Integration:**
  - Embedding AI algorithms within EHR systems to provide actionable insights at the point of care.
  - Automating data entry and reducing clinician workload.

Collaboration between Data Scientists and Healthcare Professionals

- **Interdisciplinary Teams:**
  - Forming teams that include data scientists, clinicians, IT professionals, and administrators.
  - Collaborative efforts to ensure AI/ML tools are clinically relevant and user-friendly.
- **Continuous Feedback Loop:**
  - Establishing a feedback mechanism for clinicians to provide input on AI tool performance.
  - Iterative improvements based on real-world usage and feedback.
- **Training and Education:**
  - Training programs for healthcare professionals to effectively use AI tools.
  - Ongoing education to keep clinicians updated on the latest advancements in AI/ML.

Training and Support for Clinicians

- **Comprehensive Training Programs:**
  - Initial training sessions covering the basics of AI/ML and specific tools being implemented.
  - Hands-on workshops and simulation exercises.
- **User Manuals and Guides:**
  - Detailed documentation to assist clinicians in using AI tools.
  - Quick reference guides for troubleshooting common issues.
- **Technical Support:**
  - Dedicated support teams to assist with technical issues and answer clinician queries.
  - Regular check-ins to ensure smooth operation and address any challenges.

Benefits of AI Integration in Clinical Practice

- **Improved Diagnostic Accuracy:**
  - Enhanced ability to detect diseases and conditions with higher precision.
  - Reduction in diagnostic errors and missed diagnoses.
- **Personalized Patient Care:**
  - Tailored treatment plans based on individual patient data.
  - Better patient outcomes and increased satisfaction.
- **Operational Efficiency:**
  - Streamlined workflows and reduced administrative burden on clinicians.
  - Faster turnaround times for diagnostic tests and treatment decisions.
- **Enhanced Research and Development:**
  - Accelerated pace of medical research through AI-driven data analysis.
  - Identification of new treatment approaches and potential clinical trials.

Challenges and Solutions

- **Resistance to Change:**
  - Overcoming skepticism and resistance from clinicians through demonstration of AI benefits.
  - Involving key stakeholders in the implementation process.
- **Data Quality and Availability:**
  - Ensuring high-quality data is available for AI/ML algorithms.
  - Addressing data silos and interoperability issues.
- **Regulatory Compliance:**
  - Adhering to regulatory requirements for AI in healthcare.
  - Keeping up with evolving standards and guidelines.
- **Ethical Considerations:**
  - Ensuring patient data privacy and security.
  - Addressing biases in AI algorithms to ensure fair treatment for all patients.

Slide 10: Future Trends and Innovations

- Emerging Technologies
- Potential Impact on Precision Medicine
- Future Research Directions

Emerging Technologies

- **Explainable AI (XAI): I**ncrease transparency and interpretability of AI models.
  - Greater clinician trust in AI recommendations.
  - Improved patient understanding and acceptance of AI-driven care.
- **Federated Learning:** Train AI models on decentralized data without sharing sensitive information.
  - Enhanced data privacy and security.
  - Improved model robustness through exposure to varied data.
- **Quantum Computing:** Solve complex medical problems beyond the capabilities of classical computers.
  - Accelerated drug discovery and genomics research.
  - Enhanced predictive modeling and optimization in healthcare.
- **AI-Driven Drug Discovery:** Expedite the process of discovering new drugs.
  - Reduced time and cost for drug discovery.
  - Increased success rates in bringing new drugs to market.

Future Research Directions

- **Interdisciplinary Research:**
  - Collaboration between AI experts, clinicians, geneticists, and other healthcare professionals.
  - Focus on developing holistic AI solutions that address various aspects of patient care.
- **Ethical AI Development:**
  - Research on reducing biases in AI models and ensuring fair treatment.
  - Development of ethical guidelines and frameworks for AI in healthcare.
- **Regulatory Evolution:**
  - Adapting regulatory frameworks to keep pace with AI advancements.
  - Ensuring compliance with evolving standards and guidelines.

Summary

1. Introduction
2. Importance to Precision Medicine
3. Role of AI and ML in Precision Medicine
4. Data Acquisition and Management
5. Data Privacy, Security, and Ethical Considerations
6. AI and ML Techniques in Precision Medicine
7. Case Studies in Diagnostics
8. Personalized Treatment Plans
9. Integration with Clinical Practice
10. Future Trends and Innovations

Reference:

- https://www.nature.com/articles/s41746-022-00635-4
- https://link.springer.com/chapter/10.1007/978-3-031-16443-9_60
- https://youtube.com/playlist?list=PLxI06TjqRkodv0yIr9S206KIKSl9QzIbw&si=89l8gJn-Xsljt-Fv

linktr.ee/nivu.me

Any Questions?
Thank You

NUNNARI LABS

AI Coimbatore