



GenoSync

Healthcare Solutions

Presented By: Anne Bezwada, Sai Ganesh Donka, G K Rama Naidu, Molly Moore, and Surya Tejaswi Yerramsetty



GenoSync - Healthcare Solutions

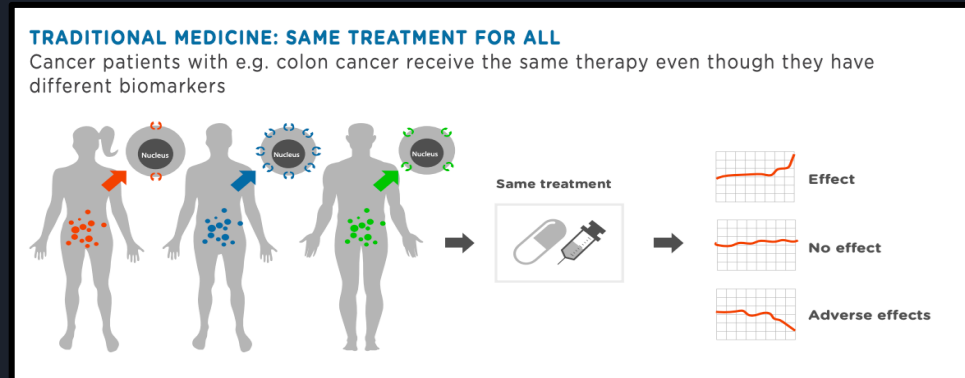


- ❖ Personalize the realm of cancer care by integrating patient **lifestyle, genomic, and clinical data**
- ❖ Promote a collaborative relationship between patients, hospitals, and healthcare teams.
- ❖ Centralized third-party data system devoted to finding patterns in aggregated patient data



Our Main Focus

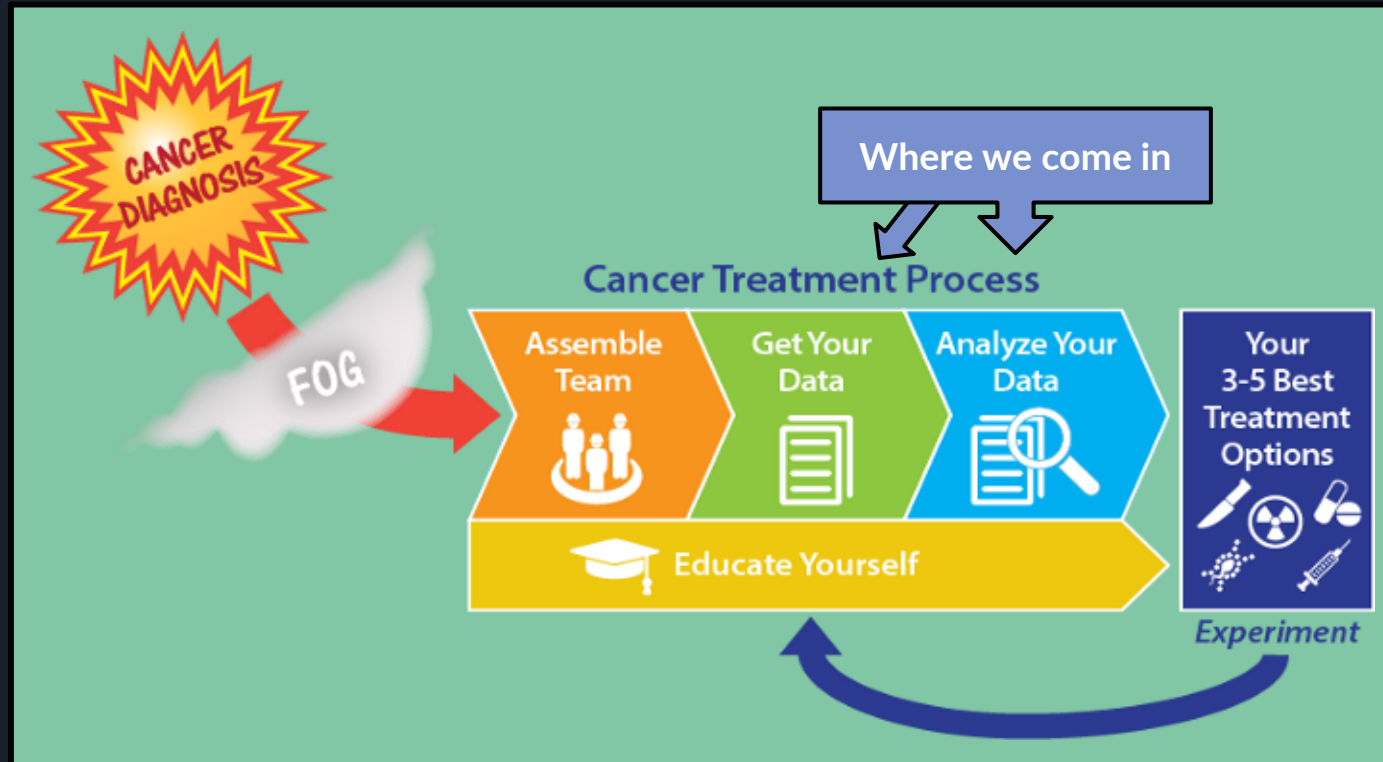
- ❖ Just over 1.9 million new cancer cases are expected to be diagnosed in the US in 2023. (American Cancer Society, 2023)
- ❖ Approx. 609,820 deaths from cancer are expected in the US in 2023, which is about 1,670 deaths per day. (American Cancer Society, 2023)
- ❖ Cancer mutates and behaves differently based on an individual's DNA
 - Why are we not choosing patient treatments based on their unique genetic data?



Picture 1



GenoSync - Healthcare Solutions



Picture 2



Overview:

- ❖ **Data Types**
- ❖ **Data Requirements**
 - Front End vs Back End
- ❖ **Data Collection**
 - Front End vs Back End
- ❖ **Flow of Data in the System**
- ❖ **Technical / System Requirements**
 - System Architecture
- ❖ **Data Analysis**
 - Machine Learning Algorithms
- ❖ **Reports**
- ❖ **Mockup UI Designs**
- ❖ **Costs**
- ❖ **Security and Compliance**
- ❖ **Risks and Solutions**



GENOSYNC - HEALTHCARE SOLUTIONS



Overview of Data Types - *Big Data*

- ❖ Our data is **High Volume** and **Highly Variable and Versatile**
 - Multiple patients with lots of historical health records
 - Many data types

- ❖ **Data Types**
 - Generic Text
 - Numerical
 - Images
 - Documents / Reports



Data Requirements - *(Front End - User Interaction)*

❖ Patient Data

- Name, age, location, generic demographic data
- Lifestyle data
 - Self-reported or collected through wearables (ex. smart watch)
- Consent forms

❖ Healthcare Data

- Patient medical records (EHR), treatment history (results)
- Clinical data
- Patient Genome data (DNA traits)





Data Requirements - *(Back End - Static Storage)*

❖ Research Data

- Drug info, efficacy, clinical trials
- Gene data - TCGA database, cancer associated genes
- Publicized cancer clinic records / treatments / research

❖ Machine Learning Algorithms

- For analyzing gene data / identifying biomarkers (*pre-built on genes data*)
- For identifying patient biomarkers in relation to available drugs/treatments (*pre-built on drug data*)

❖ Business Data

- Storage of customer data (ex. financial info, etc.)





Data Collection - *(Front End - User Interaction)*

❖ New Patient Request

- Hospital team requests to add new patient
- Consent forms sent out / retrieved

❖ Web Portal

- Allows for patient data entry
 - Lifestyle data provided by patient
 - Medical records uploaded by healthcare staff

A screenshot of the 'Patient Information' form in the GenoSync Healthcare Solutions web portal. The form is titled 'Patient Information' and includes fields for 'Today's Date' (Month, Date, Year), 'Patient's Name' (First Name, Last Name), 'Patient's Date of Birth' (Month, Date, Year), 'Email Address' (Email), 'Physician's Name' (First Name, Last Name), and 'Practice Address' (Address Line 1, City, State).

GenoSync
Healthcare Solutions

Patient Information

Today's Date

Month Date Year

Patient's Name

First Name Last Name

Patient's Date of Birth

Month Date Year

Email Address

Email

Physician's Name

First Name Last Name

Practice Address

Address Line 1 City State



Data Collection - *(Back End - Static Storage)*



❖ Genes Data

- Obtain data from The Cancer Genome Atlas (TCGA)
- European Genome-Phenome Archive (EGA)

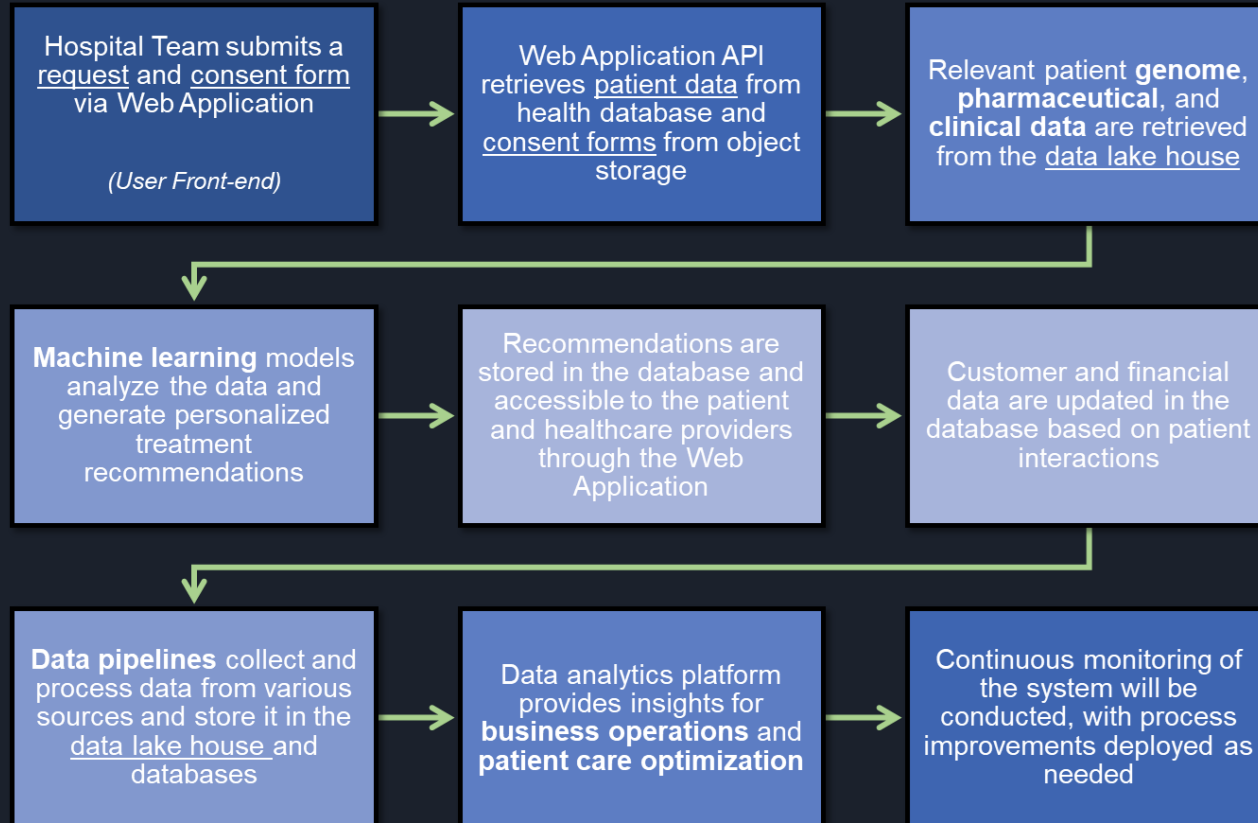
❖ Drug / Treatment Efficacy Data

- Licensed cancer drug data from National Cancer Institute (NIH)
- Cancer Drug Database (public access) from Anticancer Fund
- Clinical trial results

❖ Business Data

- Obtained from customers during initial partnership/sign-up

Flow of Data in the System





Technical / System Requirements - *(Front End)*

❖ Cloud-Based

- Web application hosted on **scalable** and **reliable** cloud service
 - AWS Amazon EC2 Service
 - S3 Std Storage
 - EBS Volumes (*for frequently updated data*)
 - Amazon RDS - secure database for patient and financial data



❖ Initial Release

- Collaborate with hospitals located in a specific region
 - Able to support 5,000 patients - growth to occur once reliability is ensured



Technical / System Requirements - *(Back End)*

❖ Cloud-Based

- Machine Learning Algorithms
 - AWS ML - Amazon SageMaker (2 instances - biomarkers/drug prediction)
- Amazon HealthLake - store clinical data (images/reports)
 - Data lake house used for unstructured data
 - AWS Lake Formation
- S3 Std Storage - store genes data and pharmaceutical / treatment data
- Amazon HealthOmics - for analysis of genomic data
- Amazon Athena - allows for Python / SQL querying

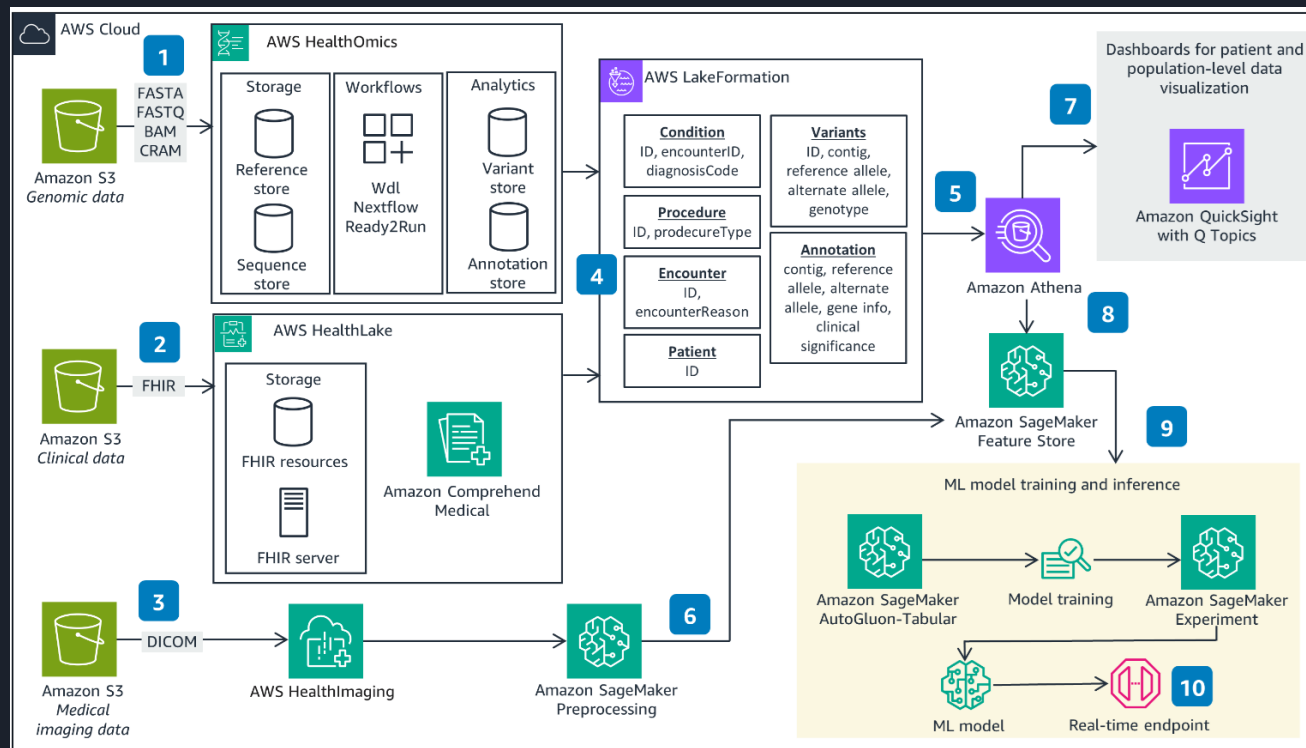




System Architecture Diagram - (Back End)

Sample AWS Architecture (right)

- ❖ For Drug data, similar architecture is used. (Amazon s3 to store and SageMaker for model training)



"Guidance for Multi-Modal Data Analysis with AWS Health and ML Services"

(Amazon - AWS, 2023)



Data Analysis

❖ Exploratory Analytics

- Pre-built Machine Learning models:
 - One machine learning algorithm focused on **genetic data - identifying biomarkers**
 - Random Forests
 - Support Vector Machines (SVMs)
 - Deep Neural Networks (DNNs)
 - One machine learning algorithm focused on **pharmaceutical / drug efficacy**
 - Convolutional Neural Networks (CNNs)
 - Recurrent Neural Networks (RNNs)
 - k-nearest neighbors (KNN)
- Work in tandem to produce a result of **best drug for identified biomarkers in patient**



Data Analysis

❖ Joint Predictions / Treatment Recommendations

➤ We aim to further develop a mechanism to **combine the predictions of both models**:

■ Example:

- **Identify top biomarkers** for Patient A with first model.
- Use second model to **predict efficacy of different drugs**, based on the identified biomarkers.
- Prioritize recommended drugs based on predicted efficacy and potential side effects.

(All while considering additional patient characteristics as an overall filter when recommending/excluding drug and treatment options)



Reports

❖ Available Reports - Early Stages

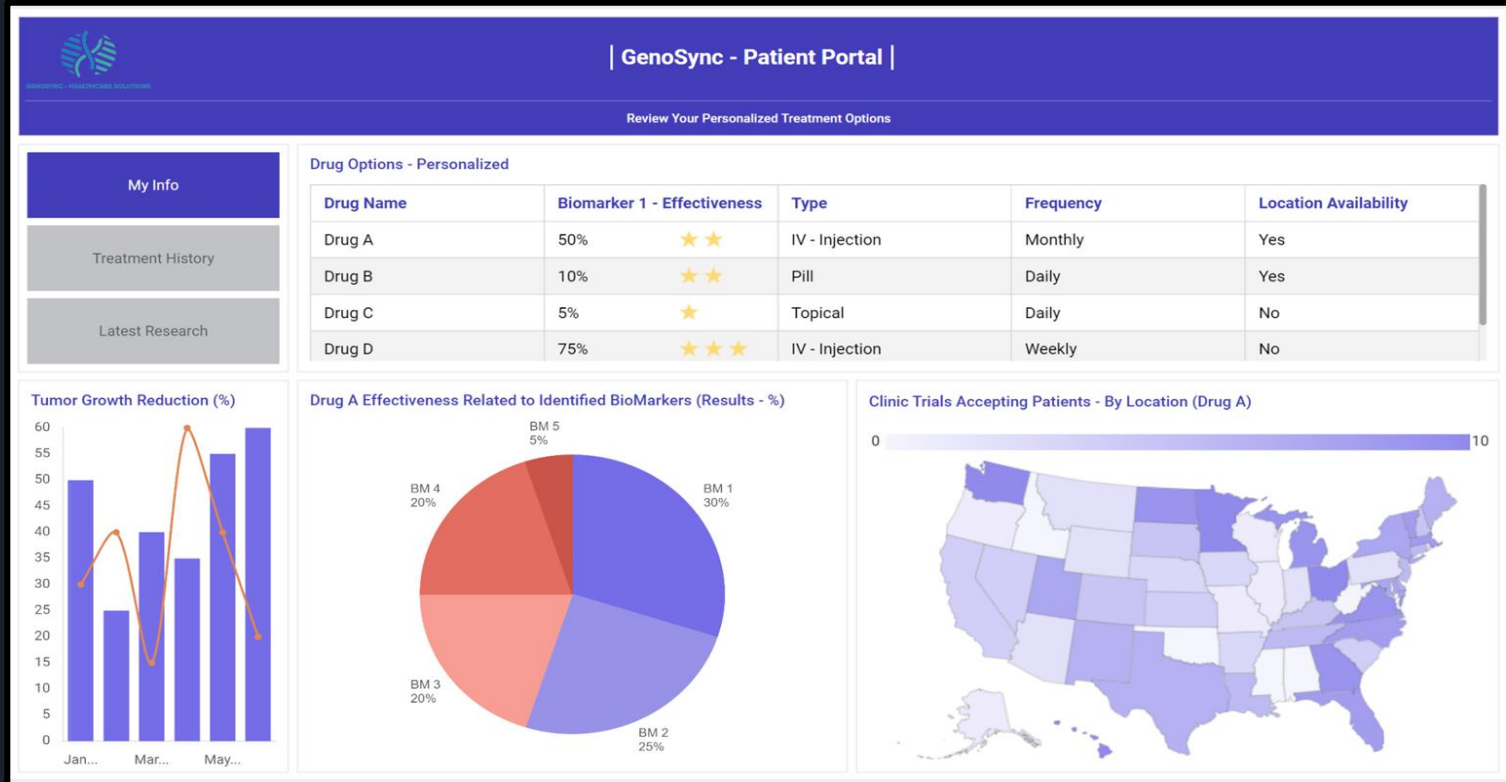
- A small selection of pre-made reports are available to healthcare providers / patients
 - Via personal dashboards
- Ensure reliability of information

❖ Available Reports - Growth Stage

- Addition of personalized reports
- Creation of simple graphical UI to allow for interactive querying of data

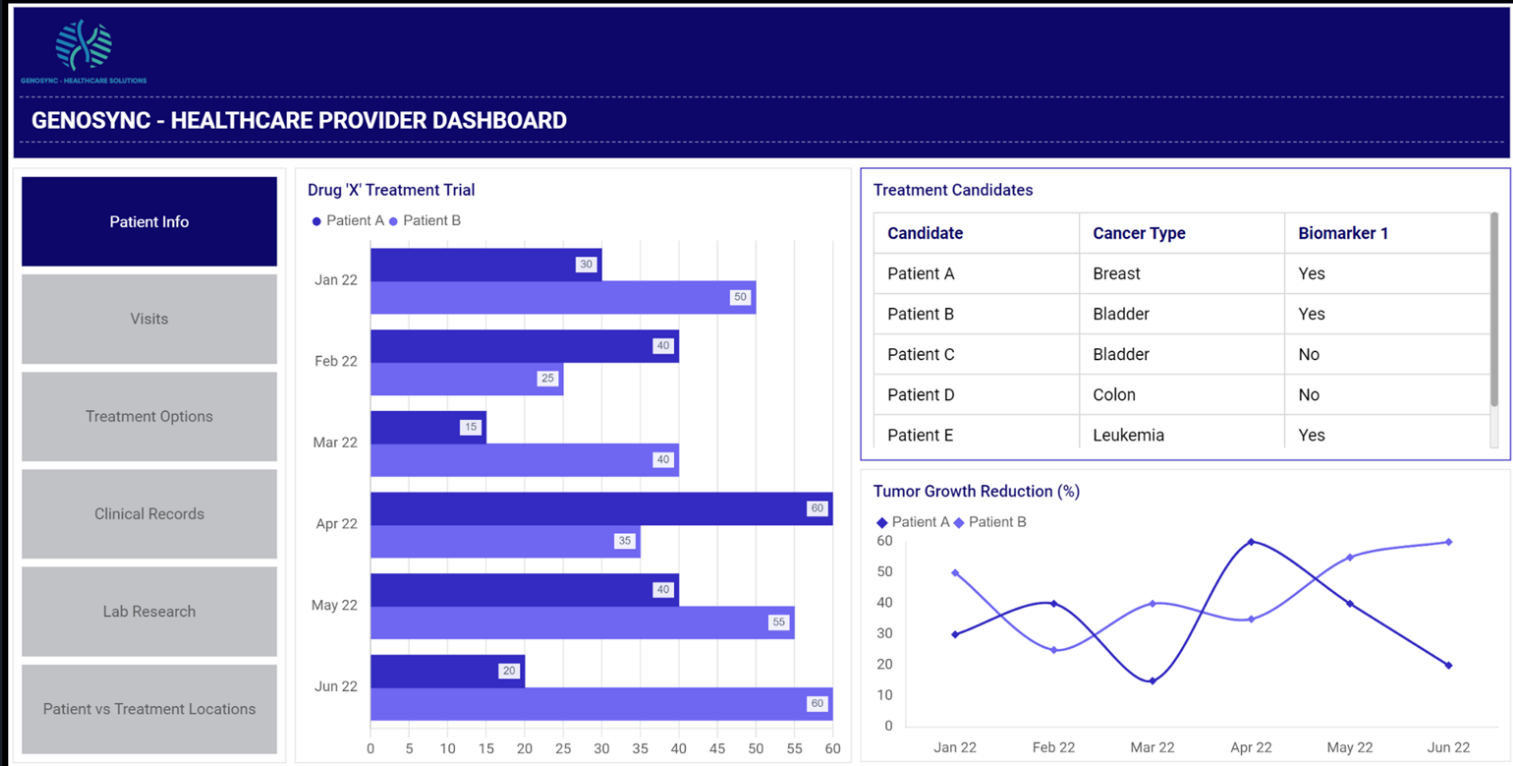


Example UI Mockup - Patient Dashboard





Example UI Mockup - Healthcare Provider Dashboard





Technical / System Requirements

❖ Staffing



- Initial staffing requirements to include:
 - 1 Cloud Architecture Engineer - to help with initial setup / structuring
 - 1 Machine Learning Developer - to set up the (2) Machine Learning Algorithms
 - 1 Full stack Developer - front end Web portal development

- Future staffing:
 - 1 - 2 Database Engineers - to provide general maintenance / management
 - Our speed of growth will determine whether we require additional staff

- **Optional:** Hire 1 -2 Sales / Marketing Reps to build customer/user relationships



Costs

Services	Components	Service Price (annual)
Amazon EC2 Service	US East	37,725.84
S3 Std Storage	US East - 4 instances	16,957.44
EBS Volumes	US East	215.64
Amazon RDS	US East	52,287.60
Amazon SageMaker	US East	21,150.72
Amazon HealthLake	US East	2,787.60
AWS LakeFormation	US East	54.00
AWS HealthImaging	US East	601.34
Amazon HealthOmics	US East	665.76
Amazon Athena	US East	1,782.48
Cloud Architecture Engineer	1 engineer, for initial development	130,000.00
Machine Learning Developer	1 developer, for initial development	50,000.00
1 Full Stack Developer	1 developer, for initial development	115,000.00
2 Database Engineers	2 engineers, maintenance/monitoring	220,000.00
Total Annual	-	649,228.42
Total (5 years)	-	1,846,142.10

(Costs estimated using AWS Price Estimator)



Security and Compliance

- ❖ Highly sensitive / regulated personal and medical data
- ❖ Compliance with federal laws (HIPAA)
- ❖ **Our Safeguards:**
 - Data encryption
 - Database Security Manager + Recovery Manager (*built into cloud-based system*)
 - Audit trails
 - Plans to address possible breaches
 - Periodic (scheduled) database backups





Risks and Solutions

❖ Precision Medicine - Early Stages

- There may not be enough clients prepared to support our data system
- Barrier to entry - need to establish trust

❖ Integration Challenges

- Data is coming from multiple sources in multiple formats
- Select proper storage formats from cloud services

❖ Clinical Decisions *(Treatments / Ethics)*

- System recommended treatments must still be reviewed/approved by medical providers
- Periodically review algorithms for any developing biases
- Deploy test environments before integrating new features





"The revolution in cancer research can be summed up in a single sentence: cancer is, in essence, a genetic disease."

Bert Vogelstein, MD

Professor of Oncology at Johns Hopkins Medical Center



Citations

American Cancer Society. (n.d.). *Cancer facts & figures 2023*. American Cancer Society.

<https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2023-cancer-facts-figures.html>

Ismail, Leila, et al. "Requirements of Health Data Management Systems for Biomedical Care and Research: Scoping Review." *Journal of Medical Internet Research*, U.S. National Library of Medicine, 7 July 2020, [www.ncbi.nlm.nih.gov/pmc/articles/PMC7380987/#:~:text=The%20requirements%20of%20a%20health,and%20\(7\)%20public%20insights.](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC7380987/#:~:text=The%20requirements%20of%20a%20health,and%20(7)%20public%20insights.)

"Guidance for Multi-Modal Data Analysis with AWS Health and ML Services." Amazon - AWS, National Council on Vocational Education, 1991, aws.amazon.com/solutions/guidance/multi-modal-data-analysis-with-aws-health-and-ml-services/.

"Precision or Personalized Medicine: Precision Medicine for Cancer." Precision Medicine for Cancer | American Cancer Society, www.cancer.org/cancer/managing-cancer/treatment-types/precision-medicine.html. Accessed 14 Dec. 2023.

Diagram Links:

Link for picture 1: <https://www.aacr.org/blog/2015/02/25/what-is-precision-medicine/>

Link for picture 2: <https://www.efpia.eu/about-medicines/development-of-medicines/precision-medicine/>