Personalized Medicine through AI and Data Analytics

Premanand S

Assistant Professor
School of Electronics Engineering
Vellore Institute of Technology, Chennai Campus

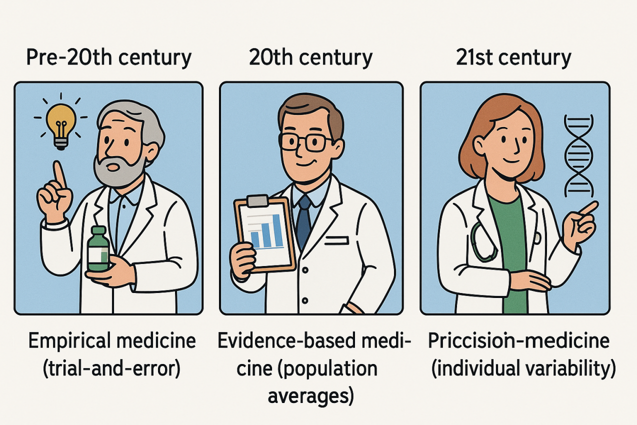
premanand.s@vit.ac.in

June 11, 2025

How many of you know
someone who tried
3+ medications before one worked?

Personalized Medicine

- A medical approach that tailors treatment and prevention strategies to the individual characteristics of each patient



Key Transition ?

- From population-level protocols to individual-level decisions
- Example: Instead of giving every cancer patient chemotherapy, use genetic profiling to see who actually benefits

Why "One-Size-Fits-All" Medicine Is Failing

- Current Limitations
  - Many drugs work for only 30–50% of the population
  - Disease progression varies dramatically between individuals
  - Side effects can be severe when treatment isn't aligned to genetic/metabolic profiles
- Examples:
  - Asthma Inhalers: Some populations don't respond due to genetic polymorphisms

Key Personalization Factors

- Genomics
  - DNA sequencing to identify SNPs, mutations, and gene expression profiles
  - Used in: Cancer therapy (e.g., HER2-positive breast cancer $\rightarrow$ Trastuzumab)
  - Used in: Pharmacogenomics (e.g., clopidogrel response by CYP2C19)
- Lifestyle
  - Diet, sleep, exercise, smoking, alcohol
  - Wearables and apps provide continuous lifestyle data
- Environment
  - Pollution, access to healthcare, socioeconomic status
  - Epigenetic changes (e.g., stress-induced gene expression modifications)

How AI is related to Personalized Medicine?

- Trial-and-error approach exists – for averages
- What if we could predict — even before starting treatment?

"Think of AI as the GPS for your health journey — it doesn't just give one route for everyone, it recalculates and personalizes based on your road, traffic, and destination."

Types of Healthcare Data

Data Type	Description	Example Use Cases
Genomic Data	DNA/RNA sequences, mutations	Cancer mutation profil-fing, pharmacogenomic
Electronic Health Records (EHR)	Clinical notes, vitals, medications, diagnoses	Risk prediction, decision support
Medical Imaging	X-rays, MRI, CT scans, Ultrasound	Tumor detection, fracture classification
Wearables & Demographics	Age, gender, occupation, habits	Social determinants of health Risk stratification

Data Challenges in Healthcare

- Technical Challenges
  - Missing values (e.g., lab tests not done)
  - Data heterogeneity (structured + unstructured text + images)
  - Temporal inconsistencies (time series, clinical timelines)
  - Integration difficulties (EHRs from multiple hospitals)
- Ethical & Legal Challenges
  - HIPAA (USA) / DPA (India): Rules for protecting personal health information
  - De-identification vs. re-identification risks
  - Bias in data: Underrepresented groups may get poorer predictions

Medical Signals vs Medical Images (ECG as Example)

Aspect	Medical Signals (e.g., ECG)	Medical Images (e.g., X-ray, MRI)
Data type	1D time-series	2D/3D spatial
Source	Electrodes/sensors	Imaging devices
Represents	Time-based activity	Anatomical structure
Format	Voltage vs time	Pixel intensities
View method	Waveforms (P-QRS-T)	Shapes, textures
File types	CSV, EDF, MAT	DICOM, PNG, JPEG
Noise	High (motion, powerline)	Medium (motion artifacts)
AI models	1D-CNN, RNN, Transformer	2D-CNN, ViT

Why Medical Imaging Needs AI ?

- Medical imaging (e.g., X-ray, CT, MRI, Ultrasound) produces large, complex visual data
  - Automating diagnosis
  - Reducing inter-observer variability
  - Improving speed and accuracy
  - Quantifying tissue features invisible to the naked eye

The quality of AI models depends on the quality and type of data.
In healthcare, data is diverse, sensitive, and deeply personal

Imagine training an AI model on data from one private hospital chain only — how well will it generalize to rural clinics?

ML vs DL in Medical Image Analysis

Aspect	**Machine Learning (ML)**	**Deep Learning (DL)**
Feature Extraction	Manual (engineered features)	Automatic (learned from data)
Input Data	Structured features (e.g., texture, shape)	Raw image pixels
Model Examples	SVM, k-NN, Random Forest	CNN, ResNet, Vision Transformer
Data Requirement	Works with small datasets	Needs large labeled datasets
Interpretability	Easier to interpret	Often a black box
Performance	Depends on feature quality	High with sufficient data
Use Cases	Pre-screening, simple classification	Tumor detection, segmentation, diagnosis

Visualization Tools for Medical Images

Tool	Use Case	Notes
3D Slicer	View, segment, and analyze MRI/CT scans	Open-source; supports Python scripting
ITK-SNAP	Manual and semi-automatic segmentation	Useful for training dataset generation
Horos / OsiriX	DICOM image viewer for macOS	Radiology workflow tool
MITK	Interactive medical image toolkit	C++/Qt; integrates with AI workflows
Napari	Large image visualization (pathology, microscopy)	Python-based, interactive, plugin support
Matplotlib / Seaborn	Overlays, histograms, and heatmaps	Used for feature maps and attention visualization

Visualization Tools for Medical Signals (e.g., ECG)

Tool	Use Case	Notes
MATLAB	Signal filtering, FFT, time-frequency analysis	Strong toolbox support for biomedical signals
WFDB Toolkit	Viewing and analyzing ECG/PPG/EEG	Comes from PhysioNet; Python and CLI versions
MNE-Python	EEG and MEG visualization	Can be adapted for ECG/EMG signals
Plotly / Bokeh	Interactive signal plots in note-books/web apps	Ideal for dashboards and presentations
TensorBoard	Visualize training loss, metrics, em-beddings	Useful during deep learning model training

Model Interpretability Tools in Medical AI

Tool/Technique	Use Case	Notes
Grad-CAM	Visualize important image regions for CNN predictions	Generates heatmaps over input images
Saliency Maps	Highlights which pixels most influence predictions	Useful for understanding CNN decisions
t-SNE	Visualize high-dimensional features in 2D/3D	Often used for clustering or class separation
PCA	Dimensionality reduction for features	Simple, fast, interpretable alternative to t-SNE
Attention Maps	Interpret attention heads in Transformer/ViT models	Shows where the model is focusing
SHAP / LIME	Feature importance for tabular and signal data	Local explanations, model-agnostic

Key Challenges in Medical AI

- Data annotation is expensive (needs radiologists)
- Bias and generalizability across different scanners/hospitals
- Regulation & ethics: AI as a medical device (e.g., FDA approvals)
- Interpretability: Trust in "black box" decisions

Practical Implementation

- Image Classification - Chest X-Ray
- Signal to Image (Scalogram)
- Segmentation

Segmentation

(a) Image

(b) Semantic Segmentation

(c) Instance Segmentation

(d) Panoptic Segmentation

Future and Research Scope

- Digital Twin
- Explainable AI
- Federated AI

Discussion

- Should Radiologists / Microbiologist fear AI or embrace it as a tool?
- How can AI support clinical decision-making without replacing humans?
- What safeguards are needed to deploy AI safely in hospitals?

Official Mail ID: premanand.s@vit.ac.in
Personal Mail ID: er.anandprem@gmail.com
Ring me: +91 73586 79961
LinkedIN: Premanand S
YouTube: ByteClassroom
Analyics Vidhya Premanand17

Healthcare data is not just big — it's complex, personal, and powerful. To unlock its value, we need both ethical responsibility and analytical creativity