Journal for Re Attach Therapy and Developmental Diversities
eISSN: 2589-7799
2023 July; 6 (7s): 1435 to 1447

Improving Breast Cancer Diagnosis and Precision Treatment with Multi-Omics Integration and Artificial Intelligence

[1*] Shruti Sharma,

[1*]Department of Biophysics, Postgraduate Institute of Medical Education and Research, Chandigarh, India, email: Shrutisharma.8722@gmail.com

***Corresponding author:** Shruti Sharma
*Department of Biophysics, Postgraduate Institute of Medical Education and Research, Chandigarh, India

Abstract

The most common malignancy in women is breast cancer (BC), which frequently poses a number of several challenges in molecular subtypes identification and early diagnosis. Additionally, approximately one-third of cases are detected at an advanced stage, owing to the fact that early-stage cancer is frequently asymptomatic, and there are presently no efficient early detection tools or particular biomarkers. This directly causes patients to miss the optimal treatment window. Unfortunately, breast cancer is highly heterogeneous and characterized by substantial heterogeneity and its various molecular typing have a direct impact on the effectiveness of treatments such as chemotherapy and immunotherapy as well as having a strong correlation with patients' rates of five-year survival rates. The aim of this review is to show how these skills help to identify novel medications and build individualized therapeutics. By creating early diagnostic models utilizing multi-omics data, identifying molecular subtypes of breast cancer, measuring the tumor immune microenvironment, and forecasting immunotherapy responses, it illustrates the successful use of AI technology in breast cancer diagnosis. These breakthroughs rely heavily on the discovery of new medications and the creation of specialized treatment regimens. The study also looks into the development of novel therapeutic agents and the mechanisms of drug resistance in breast cancer. This review examines recent developments in multimodal AI, which combines many imaging modalities such digital pathology, digital breast tomography (DBT), MRI, ultrasound, mammography, and multi-omics. Accurately identifying cancer subtypes, evaluating the tumor immune milieu, predicting immunotherapy responses, and assessing possible treatment resistance are all made possible by this integration. Artificial intelligence (AI) has progressed rapidly from experimental research to therapeutic applications. It can now achieve diagnosis accuracy equivalent to that of radiologists, improve specificity, and reduce workloads by 44% to 68%, all while maintaining cancer detection rates. Some research indicates that artificial intelligence might help radiologists identify more cancer patients. These advantages result in more effective screening procedures, less needless recalls, and quicker diagnosis. In the future, artificial intelligence technology is likely to give more accurate and individualized diagnostic and therapy alternatives for patients with BC. However, analyzing multi-omics data in cancer involves a number of problems, including complexity, sparsity, transparency, and ethical concerns.

KEYWORDS: Breast cancer, AI, omics, personalized medicine

Introduction:

A major contributor to cancer-related deaths among women globally is breast cancer. Timely and accurate diagnosis is crucial, as clinical outcomes can be significantly improved with early and precise detection. In order to provide a thorough understanding of biological processes, identifying disease pathways and uncovering novel therapeutic approaches, multi-omics analysis synthesizes many data sources. A more thorough understanding of molecular mechanisms, disease pathways, and possible treatment. Multi-omics approaches integrate heterogeneous datasets to facilitate a comprehensive understanding of biological processes, reveal disease pathways, and inform the discovery of novel therapeutic strategies. While each molecular layer provides a unique viewpoint; however, when examined collectively, the comprehension of biological functions, signaling cascades, and expression patterns is greatly improved by these layers. However, combining these layers presents considerable difficulties, but the combination of these layers also introduces significant analytical challenges. Advanced methodologies are needed to extract relevant insights since different platforms have different data formats, measurement scales, and dependability **(**Js et al., 2023).
By improving the methods of disease diagnosis, prediction, and treatment, artificial intelligence (AI), including machine learning (ML) and deep learning (DL), is transforming the healthcare industry **(**Lu et al., 2021)**.** Large-scale analysis is a strength of these computational tools, which often uncover links and patterns that traditional statistical methods miss when analyzing omics information in several dimensions. Multidimensional omics datasets are effectively analyzed using these computational methods, which can reveal complex connections and patterns, which are often overlooked by

traditional statistical approaches. Diagnostic accuracy improved in terms of identification of novel targets and potential for specific treatment strategies. The outcome is increased in diagnostic accuracy, the identification of new targets, and the potential for tailored treatment approaches**.** Not with standing these improvements, maintaining interpretability is difficult. Although, these developments, there are still issues to be resolved, such as the requirement for large annotated cohorts and the harmonization of various data sources and the difficulty of preserving interpretability. In order to responsibly and successfully implement AI in medical research, it is imperative essential to address these issues (I.H et al., 2021).The integration of multi-omics using AI-driven methods has excellent great potential for improving individualized therapy. AI systems are more effective better at discovering complicated linkages, forecasting illness patterns and identifying biomarkers when they analyse complex, high-dimensional data. In comparison to traditional procedures, this is made possible by the integration of many biological data types. However, the quality of the data, the intricacy of the analytical work and the thoroughness of the validation process often affect, information security, potential algorithmic biases, and patient data privacy are some of the elements that determine the study's success. Developments in this subject raise important ethical issues. Establishing precise rules and standards on these matters is crucial to securing patients' interests and ensuring moral behaviour. To ensure responsible application and protect patient interests, clear regulations and standardized guidelines must be established. By utilizing various biological data types, AI systems may find complicated associations, anticipate illness trends and identify biomarkers more efficiently than many traditional methods when working with complex, high-dimensional data. Nevertheless, the effectiveness of the analysis is frequently determined by the quality of the data, the complexity of the analytical work and the thoroughness of the validation process. The developments in this sector create ethical questions concerning patient data privacy, information security, and algorithmic prejudice. To guarantee moral usage and protect patient interests, it is essential to establish clear rules and regulatory standards. The incorporation of modern clinical inputs and the application of transparent AI techniques to improve model interpretability can greatly increase trust and encourage wider implementation in healthcare settings. A wider implementation of these models in healthcare settings can be facilitated by utilizing interpretable AI techniques and integrating recent clinical discoveries. According to our research, the majority of AI-based multi-omics systems are still in the proof-of-concept phase and have little use in actual clinical situations. Despite their potential in areas such as survival prediction and subtype classification, these methodologies frequently need substantial external validation, practical clinical research, and regulatory approval. A strong strategy for enhancing improving our understanding knowledge and ability to treat diseases is the combination of artificial intelligence with multi-omics data analysis. In the healthcare industry, this synergy increases the possibility of creative solutions and improved outcomes. Finally, combining AI with multi-omics data analysis provides a strong tool for better diseases knowledge and management. By overcoming technological constraints, enhancing model transparency, and proactively addressing ethical hazards, the profession may fully utilize these tools. We can make the most of these technologies by addressing these technological issues, improving model clarity, and addressing ethical issues. Ultimately, this will lead to more accessible, individualized, and targeted healthcare, which will help advance precision medicine (Ching al., 2018; Subramanian et al., 2020).

Transcriptomic, proteomic, lipidomics and metabolomic techniques

Understanding the transition from genotypic effects to genomic data is crucial for comprehending disease progression and treatment impacts in breast cancer, where omics levels describe the phenotypic traits. Bridging the gap from genotypic effect to genomic data is essential for gaining a thorough understanding of disease progression and medication effects in breast cancer and omics levels, are used to describe the phenotypic traits bridging the gap between genotypic effects and genomic data is crucial for a comprehensive understanding of disease progression and drug responses in breast cancer, while omics analyses are employed to characterize phenotypic traits (Hasin et al., 2017; Kristensen et al., 2014;)**.**

Transcriptomics:

The term "transcriptome" refers to all of the mRNA in a subject or sample. Two contemporary high-throughput sequencing methods used to obtain transcriptomic data are RNA sequencing (RNA-Seq) and microarrays. mRNA expression is determined by microarray analysis by quantifying the hybridization between complementary probes and samples. The fluorescence intensity seen in each array well serves as a gauge for the sample's level of gene expression. The degree of gene expression in the sample is represented by the fluorescence intensity in each well of the array. The fluorescence intensity in each array well indicates the level of gene expression in the sample or utilization of probe in contrast to microarray analysis. Both quantifying mRNA expression levels and discovering novel sequences. Microarray analysis is limited because the gene sequence must be known in order to design probes, making it useful for both quantifying mRNA expression levels and identifying novel sequences. The method being used is similar to Sanger sequencing, involves sequentially introducing fluorescently tagged nucleotide bases to determine the mRNA sequence. The precise sequence and level of expression are revealed by analysing fluorescent images captured during each iteration. The labour needed for microarray analysis is less. However, RNA-Seq needs sample preparation and can undergo with fewer samples, while RNA-Seq25 does not require previous knowledge of gene sequences. Both methods are capable of

high throughput, but microarrays currently have a greater cost-value (Cheng et al., 2025; Kong et al., 2021). Analysing the frequent RNA-Seq or microarray analysis and characterize different disease phenotypes and pharmacological effects within a system, giving important data for the development of genome-specific therapies (Hagiwara et al., 2023), RNA-Seq presents a valuable opportunity for the identification of new genomic drug effects. Identification new genomic drug effects presents a valuable opportunity of RNA-Seq and illness phenotypes. Microarray assays, on the other hand, are more standardized and less costly. Because RNA-Seq has a greater signal-to-noise ratio than microarray data, it is frequently chosen for clinical research. Because RNA-Seq can generate a greater signal-to-noise ratio than microarray data, it is often chosen for clinical investigations for medical studies. Additionally, RNA-Seq requires significantly less sample material than microarray approaches with nanogram and microgram amounts, respectively (Hagiwara et al., 2023; Kong et al., 2021). It is anticipated that RNA-Seq procedures will become more standardized and take the place of microarray diagnostics as NGS become more integrated into clinical diagnostics. Based on financial and experimental needs, both diagnostic approaches are now utilized to generate transcriptomic results (Hagiwara et al., 2023). As shown in Figure 1.

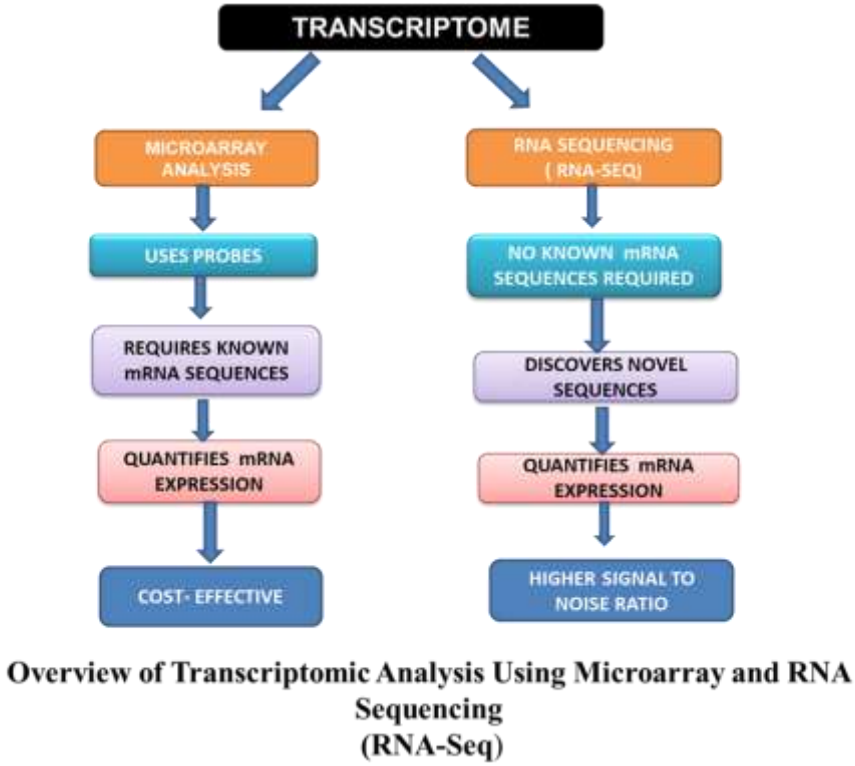

Figure 1: Transcriptome analysis methods: A comparative overview of microarray and RNA-Seq workflows, methodologies, and applications in gene expression profiling.

NGS

Correlating imaging phenotypes with gene expression is made possible by the incorporation of AI-assisted radiogenomics. This method makes it easier to classify patients for targeted treatments and identify the genes that most significantly interact with BC. From histopathological images, spatially resolved gene expression can be predicted using convolutional neural network (CNN) models. Without the need for expensive molecular tests, this approach makes it easier to identify over 100 genes, A technique called hist2RNA was created by Mondol et al. to predict the expression levels of 138 genes using full-slide photographs. With a correlation of 0.82 between patients and 0.29 between genes, this approach effectively predicted gene expression in a held-out test set (n = 160) (Mondol et al., 2025). In order to forecast 250 genes, Rahaman et al. assessed four distinct cutting-edge deep learning architectures, none of which used pre-trained weights. A median correlation coefficient greater than 0.50 was found for 24 of the 237 positively linked genes. High-dimensional gene expression patterns in tumor RNA sequences may be analysed by artificial intelligence to determine the stages and functional characteristics of BC (Rahaman et al., 2025). One method uses whole-slide photos to predict the expression levels of 138 genes. The approach effectively identified the targeted genes in a held-out test set of 160 samples, achieving a correlation of 0.82 across patients and 0.29 across genes (Mondol et al., 2023). High-dimensional gene expression patterns in tumor RNA sequences may be analysed by AI to determine functional traits and stages of BC, which may provide important biomarkers for various BC stages for future study. Using deep learning techniques, Kim et al.

investigated the connection between gene expression and BC metastasis, revealing an area under the curve (AUC) of 0.754 **(**Kim et al., 2023). In a related work, Jung et al. used XGBoost models in conjunction with gene expression patterns to identify metastatic marker genes (MGs) associated with BC in order to reveal functional traits and stages of BC **(**Jung et al., 2023**).** Overall, AI has the ability to evaluate high-dimensional gene expression patterns in tumor RNA sequences, resulting in the identification of key biomarkers for various stages of cancer for future research. After employing deep learning techniques to investigate the relationship between gene expression and BC metastasis, Kim et al. found an area under the curve (AUC) of 0.754 **(**Li et al., 2022**).** In a separate work, Jung et al. identified metastatic marker genes (MGs) important to BC using XGBoost models in conjunction with gene expression patterns. Three MGs-SPPL2C, KRT23, and RGS7-showed statistically significant results in survival analysis, with a P-value of less than 0.01, according to their research **(**Jung et al., 2023; Kim et al., 2023). Researchers can use a variety of methodologies, such as mutual information-based analysis, correlation analysis, gene co-expression networks, and feature selection techniques within machine learning (ML) and deep learning (DL), to find possible biomarkers. In their search for BC metastatic marker genes, Jung et al. also discovered an AUC of 0.754. Gene co-expression networks, mutual information analysis, correlation analysis, and feature selection in machine learning and deep learning may all be used to find potential biomarkers **(**Abidalkareem et al., 2024).

Microarray:

To evaluate the potential for biomarker molecular expression in malignant tissue, use only the tissue architecture displayed in digitized samples stained with hematoxylin and eosin (H&E). The study evaluates the possibility of molecular expression of biomarkers in cancer tissues using tissue architecture shown in digitized hematoxylin and eosin (H&E) stained samples. The BC tissue microarray collection from patients at Vancouver General Hospital in British Columbia, Canada, was the subject of the retrospective diagnostic study, which was conducted at a single location **(**Lu et al., 2021**).** The investigation and analysis was place between July 1, 2015, and July 1, 2018, where the inquiry and analysis took place. To evaluate the potential for biomarker molecular expression in malignant tissue, use only the tissue architecture displayed in digitized samples stained with hematoxylin and eosin (H&E) **(**Coudray et al., 2018). Morphological-based molecular profiling (MBMP) is a machine learning approach. The biomarker expression in the tissues under study was predicted using a deep convolutional neural network and the correlations between biomarker expression and histomorphology was investigated using logistic regression. From July 1, 2015, to July 1, 2018, the study and analysis were carried out. 20,600 digitized, publicly available H&E-stained sections from 5,356 BC patients from two cohorts are included in the collection. Two cohorts-412 patients in cohort 1 and 4,944 individuals in cohort 2-were used for the investigation and analysis. Morphological-based molecular profiling (MBMP), a machine learning technique, was created. This technique predicted biomarker expression in the tissues under study using a deep convolutional neural network. Furthermore, logistic regression was used to examine the connections between biomarker expression and histo-morphology. The median age of diagnosis in cohorts 1 and 2 was 61 and 62 years, respectively. 20,600 digitized, publicly available H&E-stained sections from 5,356 patients with BC from both cohorts were included in the dataset. Progesterone receptor (PR), estrogen receptor (ER), and ERBB2 (formerly known as HER2) were among the 19 biomarkers whose molecular expression was shown to be significantly correlated with tissue histomorphology, the results were determined to be non-inferior to traditional immune-histochemistry (IHC). Diagnoses becoming more precise as more information becomes accessible, A Bhattacharyya distance of 0.25 suggested that patients with an ER-negative/PR-positive IHC status had substantially different morphology than patients with an ER-negative/PR-negative IHC status **(**Jiang Y et al., 2022**).** With a Bhattacharyya distance of just 0.03, the morphology of patients with an ER-negative/PR-positive state was comparable to the patients with an ER-positive condition. Because the IHC results are occasionally misreported as negative, these individuals may need anti-hormonal medicine. The study findings, which demonstrated a Positive Predictive Value (PPV) of 91–98%, a Negative Predictive Value (NPV) of 51–78%, and an overall accuracy ranging from 81% to 90%, were in line with traditional IHC techniques. MBMP predicted both IHC and biomarker expression in at least half of the trial participants. At least in half of the trial participants, both IHC and biomarker expression was predicted by MBMP. The results show that training data improves prediction accuracy, morphological-based molecular profiling, which uses digitalized H&E-stained images, is a common method for large-scale molecular profiling. This method makes it possible to profile several biomarkers in cancer tissues at once quickly, accurately, and affordably. The findings suggest that as the amount of data used for training increases, prediction accuracy will likely increase, as shown in Figure 2 **(**Jiang et al., 2022; Kather et al.,2020**).**

Journal for Re Attach Therapy and Developmental Diversities
eISSN: 2589-7799
2023 July; 6 (7s): 1435 to 1447

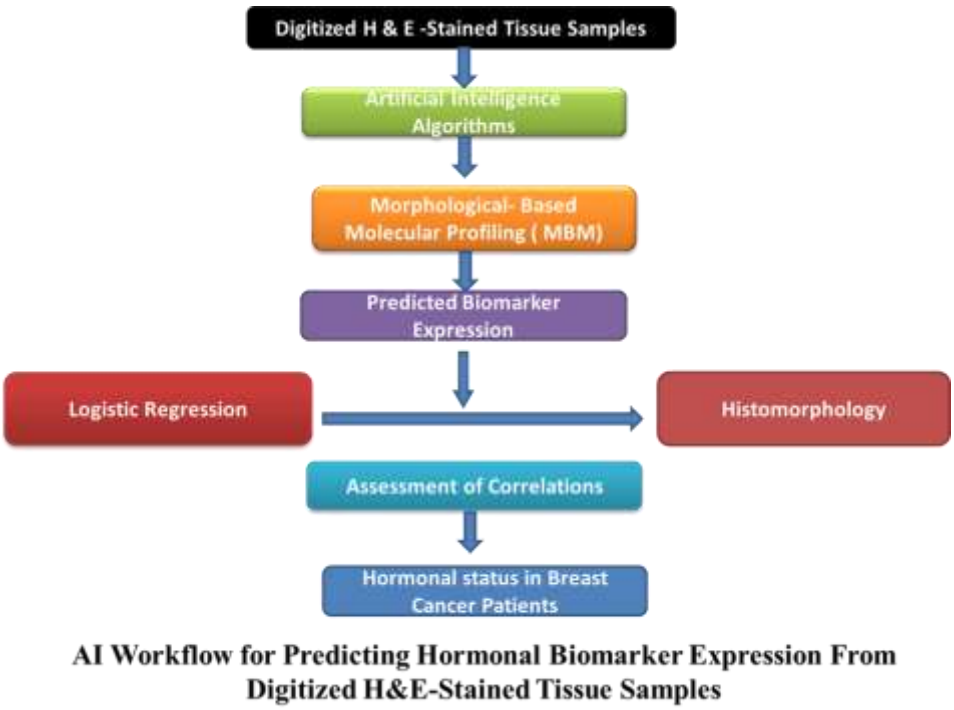

Figure 2: Artificial Intelligence–Driven workflow for predicting hormonal biomarker expression from digitized hematoxylin and eosin (H&E)–Stained Tissue Samples through computational histopathological image analysis.

Breast cancer microarray data is classified into normal and relapse using simple machine learning techniques. The gene expression omnibus (GEO) website, specifically GSE45255 and GSE15852, provides the information. To create a single dataset, these two datasets are integrated and combined. The study utilized three machine learning algorithms: random forest (RF), extra tree (ET), and support vector machine (SVM). The algorithms' hyperparameters were optimized using grid search cross validation (CV). According to the findings, the adjusted SVM has the highest accuracy among the algorithms tested, at 97.78%. It is advised to incorporate a feature selection approach in the future to identify the best features and enhance classification accuracy **(**Chaudhary et al., 2018**;** Duan et al., 2007).
To classify microarray data, Jinthanasatian et al., used a firefly algorithm with neuro-fuzzy. The parameters were optimized using a firefly algorithm, and an informative gene selection and rule set creation were performed using a neurofuzzy technique. Using 10 k-fold method, the accuracy of the suggested technique was evaluated on seven datasets. Although the suggested method yielded comparable results with other approaches, there is still room for improvement, particularly in the colon dataset, which only reached 76.94% **(**Jinthanasatian et al., 2017).To better analyze microarray data, Li et al., suggested random value-based oversampling (RVOS) and an enhanced version of SVM-RFE. First, the distribution of two samples was balanced using RVOS. Next, the informative genes were obtained using an enhanced version of linear SVM (LLSVM) with the improved RFE method. The suggested model was ultimately tested utilizing four classifiers with stratified 5-fold CV. The findings demonstrated that SVM-VSSRFE outperformed LLSVM-VSSRFE in three datasets, as well as the latter's effectiveness in lowering time use, particularly when dealing with datasets with large dimensionality **(**Li et al., 2018**).**

Microarray-Based AI Studies

Study	**Technique**	**Accuracy**	**Notes**
SVM (Grid Search CV)	GSE45255 + GSE15852	97.78%	Highest accuracy
Firefly + Neuro-fuzzy	7 datasets	~76–95%	Colon dataset lowest
RVOS + SVM-RFE	4 datasets	Improved	Balanced sample distribution

Table 1:- This table shows Microarray – Based AI studies that use artificial intelligence techniques to examine gene expression patterns and find possible biomarkers or diagnostic signs are compiled in this table.

Proteomics:
The methodical process of identifying and classifying every protein found in a biological system and clarifying the connections between these proteins is referred to as "proteomics". Important details about protein quantity, structure, and

subcellular location are obtained by proteomic investigations. Additionally, these analyses measure the rates of protein production and degradation and aid in the comprehension of protein-protein interactions. This all-encompassing method greatly improves our understanding of cellular mechanisms and promotes developments in a number of scientific and biological domains. Information from this data is used for personalised medicine (PM), as it helps us identify illness patterns and comprehend how the proteome evolves during different biological events **(**Wei et al., 2019**).** Data on post-transcriptional changes, such as phosphorylation, glycosylation, and acetylation, is utilized for PM. Changes in a tissue's protein abundance, makeup, or structure can have a big impact on the course, diagnosis and treatment of disease. Mass spectrometry (MS) has been the primary method for acquiring proteomic data for the last twenty years, notably for determining protein expression and modification sites, as well as for examining protein-protein interactions **(**Demircioglu et al., 2023)**.**Bottom-up and top-down proteomics are the two main approaches for producing proteomics data often called "shotgun proteomics," the bottom-up approach entails breaking down proteins into peptides, employs mass spectrometry to examine the composition of large, complex protein samples. The bottom-up method is often helpful for analysing an unidentified protein combination for a variety of reasons, but it also provides inaccurate data about a particular protein, data from mass spectrometry (MS) is frequently misconstrued, particularly when it is fragmented. The MS output usually only detects proteins that are present in high concentrations in the mixture, which could result in the loss of crucial information. Notwithstanding this drawback, the shotgun proteomics method is useful in PM since it makes it possible to create a distinct proteome "fingerprint" for every person. This method may assist in locating important protein markers for particular illnesses. It is now easier to examine changes in proteomics linked to biological disruptions thanks to recent labeling techniques that allow for the simultaneous analysis of several samples employing shotgun proteomics (bottom-up labeling) which may be MS data that is easily misinterpreted when broken, the MS output only shows proteins with large concentrations in the mixture, and it may be lost. Only proteins with high concentrations in the combination are visible in the MS output, and it can be lost. Nevertheless, the shotgun approach is helpful in PM since it facilitates the creation of a distinct proteomic "fingerprint" for every individual. This may lead to the discovery of important protein markers for certain illnesses. Specifically, a number of illnesses are associated with post-translational changes, including blood disorders, cancer, diabetes, infectious diseases, and neurodegenerative diseases **(**Han et al., 2012). Identifying critical post-translational modifications in certain patients may be a powerful diagnostic tool in the PM setting. Additionally, professionals can obtain thorough pharmacodynamic information on therapeutic treatments by analysing the temporal expression of a certain protein **(**Castiglioni et al., 2021**).** Finding significant post-translational changes in some patients could be a useful diagnostic technique in the PM context. A hybrid technique known as "middle-down proteomics" has developed in recent years to incorporate the benefits of both strategies. Middle-down proteomics uses protein digestion, just as bottom-down proteomics, but aims to create much bigger peptides. In recent years, a hybrid technique known as "middle-down proteomics" was developed to incorporate the benefits of both strategies. Similar to bottom-down proteomics, middle-down proteomics makes use of protein digestion, but aims to create much bigger peptides, which leads to protein solutions with fewer ambiguities and complexities, as well as the capacity to analyse high-level features. Middle-down proteomics has previously been proven to be the most effective method for researching histone proteins **(**Blanksby et al., 2010). Additionally, it has potential for PM applications since it enables the quantification of a variety of putative protein biomarkers as well as the investigation of specific protein mutations and alterations. Furthermore, it makes it possible to examine specific protein mutations and alterations as well as quantify a variety of potential protein biomarkers (Blanksby et al., 2010; Castiglioni et al., 2021; Han et al., 2012).

LCMS

This research seeks to find reliable metabolite biomarkers of BC and triple-negative breast cancer (TNBC) using a serum metabolomics methodology. The serum metabolic profile of the BC group (n = 53) was compared to that of the non-BC group (n = 57), as well as the variations between TNBC patients (n = 23) and non-TNBC participants (n = 30), using an untargeted metabolomics strategy utilizing ultra-high performance liquid chromatography coupled with mass spectrometry (UHPLC-MS) **(**Zhang et al., 2020**).** The differential metabolites were identified using multivariate data analysis, the Mann-Whitney U test, and the fold change calculation. Furthermore, logistic regression analysis and receiver operating curve analysis, among other machine learning approaches, were used to create diagnostic biomarker panels. Between the BC and non-BC groups, there were 36 metabolites that were significantly different, and between the TNBC and non-TNBC patients, there were 12 metabolites that were noticeably different. With an area under the curve (AUC) of 0.995, the findings also demonstrated that fourmetabolites-N-acetyl-D-tryptophan, 2 arachidonoylglycerol, pipecolic acid, and oxoglutaric acid were regarded as essential biomarkers for the diagnosis of BC and non-BC. Another two-metabolite panel consisting of N-acetyl-D-tryptophan and 2-arachidonoylglycerol was discovered to differentiate TNBC from non-TNBC and yielded an AUC of 0.965 **(**Budczies et al., 2013;Tang et al., 2021).

Metabolomics Biomarker Findings-

Study	Sample Size	Method	Key Biomarkers Identified	AUC / Accuracy

UHPLC-MS untargeted	53 BC vs 57 control	LC-MS	4 metabolites	AUC = 0.995
TNBC subset	23 TNBC	LC-MS	2 metabolites	AUC = 0.965
Plasma Proteomics	335 patients	MS	Classifier proteins	AUC = 0.96

Table 2:- This table shows important metabolomics biomarker discoveries are compiled which shows noteworthy compounds found by metabolomics profiling that could be useful markers for the onset, course, and response to therapy of a disease.

Mass spectrometry

Mammography is still the gold standard for screening for breast cancer, but in some patient groups-especially women with thick breast tissue-its accuracy may be low. A useful addition to conventional screening techniques is liquid biopsy testing. In order to create a plasma-based protein classifier that can differentiate between people with early-stage breast cancer and healthy people, a case-control research was carried out. The goal of this research was to develop a classifier that can accurately differentiate between those developing early-stage breast cancer as compared to those without BC (Zhang et al., 2020).. Mammography is still the gold standard for screening for breast cancer, but in some patient groups like women with thick breasts its accuracy is still low. Tests based on liquid biopsy provide a readily available complement to conventional screening techniques. In order to create a plasma-based protein classifier that can differentiate between people with early-stage breast cancer and healthy people, a case-control research was carried out, where semi-quantitative, label-free mass spectrometry (MS) in conjunction with a sample preparation technique was used to blindly evaluate plasma samples from a population of 335 women **(**Geyer PE et al. 2021)**.** This group comprised 116 patients with newly diagnosed, treatment-naive breast cancer at stages 0–2 and 219 healthy controls. The median number of proteins found per patient in the breast cancer group was 6,991, compared to 6,818 in the healthy control group **(**Johansson et al. 2019). A machine learning-based classifier was created and tested to detect people with breast cancer using patient proteome profiles and a leave-one-out cross-validation (LOOCV) method. In spite of breast density, the classifier showed a sensitivity of above 85% in individuals with breast cancer. In particular, while maintaining a specificity of 90%, the sensitivity was found to be 87.2% for individuals with low breast density and 90.2% for those with high breast density. It was discovered that the area under the receiver operating characteristic curve (AUC) was 0.96 (95% CI: 0.93-0.97)**.** The method's specificity was 90.4% (95% CI: 85.8-93.6%) and its sensitivity was 86.2% (95% CI: 78.8-91.3%). These results demonstrate plasma proteomics' enormous potential as a useful diagnostic method for breast cancer early detection **(**Blixt et al., 2022).Using patient proteome profiles and a leave-one-out cross-validation (LOOCV) approach, the effectiveness of a machine learning classifier we created to detect people with breast cancer. In a study it was shown that plasma proteomics can be an effective diagnostic tool for breast cancer early detection **(**Johansson et al., 2019; Blixt et al., 2022**).** The wide variety of cellular subtypes in breast cancer makes its molecular pathogenesis a difficult problem. Accurately and quickly diagnosing and forecasting the course of disease is considerably enhanced by the capacity to identify and see the distribution of various cell subtypes at the single-cell level inside tissue samples. In a study, the molecular profiles of 14 different breast cancer cell lines at the single-cell and subcellular levels was collected and displayed using mass spectrometry imaging. a molecular library of cell lines with strong genetic characterization was collected. The creation of a genotypic identification model based on metabolic characteristics was made easier by multistep processing, which was applied to deep learning techniques. In addition to evaluating hormonal state and breast cancer subtypes, this model showed cross-validation rates as high as 97%. Additionally, during the measurement procedure, our single-cell identification models were used on complicated tissue samples, allowing for the quick identification of cell subtypes inside a particular tissue context in a matter of seconds. A genotypic identification model based on metabolic phenotypes with cross-validation rates of up to 97% as well as the hormone status of the cancer and a subtype of breast cancer was produced by multistep processing, which incorporated deep learning. Furthermore, single-cell-based identification models to complicated tissue samples was extended, allowing to identify cell subtypes within a tissue context in seconds. These data serve as the foundation for quick and precise diagnoses and prognostics with high spatial resolution by using MSI to show digital pathology "on the spot" at the single-cell level.

Metabolomics:

Intermediate products of metabolic processes with a low molecular weight are known as metabolites. Their identification and analysis are the main goals of the field of metabolomics. Metabolomics demonstrate how an organism is impacted by both hereditary and environmental variables and metabolites are effective indicators. A "functional readout" of the organism's current state is often used to describe a thorough study of metabolites. Metabolomic data can offer vital insights about a person's distinct physiological reactions to medications in the setting of pharmacometabolomics. Because metabolites reflect both hereditary and environmental factors, they are useful. A thorough metabolic study is frequently thought of as a "functional readout" of the organic system's current condition (Baciu et al., 2021)**.** In a PM environment, metabolomic data can offer insights into a person's unique physiological reaction to a medicine, which is sometimes called

metabolomics. The discovery of biomarkers for diseases that may be treated by PM has helped in the development of PM methods, helped doctors with diagnosis, and aided in early therapy, all thanks to the current metabolomics study of tissues and biofluids. The ability of metabolomics to conduct non-invasive tests using easily accessible biofluids, such as urine and blood, since metabolites, unlike the majority of proteins, spread throughout the body and can be found in urine (Witting et al., 2015).Nuclear magnetic resonance (NMR) spectroscopy was the main technique used for metabolite detection in the early phases of metabolomics research. However, mass spectrometry (MS), which offers better resolution and increased sensitivity for identifying low amounts of metabolites has been increasingly popular during the last ten years (Duan et al., 2007).Targeted and untargeted methods are the two primary categories of proteomics, a subset of metabolomics. The metabolomic profiles of a number of illnesses, including diabetes, liver disease, Crohn's disease, and Parkinson's disease, have been identified, thanks to the untargeted technique. A metabolomic approach can be broadly classified into two categories targeted and untargeted methods. Utilizing the untargeted or global, approach for multiple malignancies and diseases (Blixt et al., 2022). It is difficult to examine the vast range of metabolite concentrations in a typical sample, which is thought to span 7–9 orders of magnitude. No single technology can generate a complete fingerprint of all metabolites, ranging in magnitude from pmol to mmol. However, improvements in MS resolution have made it possible to identify large metabolites, and developments in liquid and gas chromatography techniques, made it possible to achieve cleaner separations of metabolites and distinct peaks at various concentration levels (Baciu et al., 2022). The difficulty in correctly detecting the thousands of peaks produced by an untargeted experiment, which has a high percentage of false positives, remains this approach's greatest flaw. Even these rates present challenges for clinical adaption in PPM38. Untargeted metabolomics is an essential method for generating hypotheses about possible biomarkers (Mondol et al., 2023).Targeted metabolomics, focuses on measuring known metabolites in particular biological samples, is a significant area of metabolomics research in the field of precision medicine (PM). By employing a focused strategy, medical professionals can examine patient biofluids for aberrant metabolite concentrations that could be markers for specific diseases. Most metabolomics research in PM focuses on targeted metabolomics, which seeks to quantify known compounds in a specific sample. Using the targeted approach, clinicians may analyse a patient's bio fluids for abnormal metabolite concentrations, as diagnostics. This approach can be used by clinicians to determine the optimal dosing schedule and evaluate metabolic reactions after medication administration. However, this strategy must include the identification and complete validation of suitable metabolic biomarkers in order to have clinical relevance. When paired with other omics data, metabolomics has great potential to advance customized medicine. For this project, the application of bioinformatics techniques is essential. Strong methods for integrating omics data are beginning to emerge in the field. One of these applications is Metabox, a free R-based application that integrates MKGI models with genomic, proteomic, and metabolomic data. These models make use of neural networks to find connections between various transcriptomic and omics datasets. When combined with other omics data, metabolomics has a promising future as a tool for PM advancement (Kim et al., 2023; Abidalkareem et al., 2024). The use of bioinformatics methods is essential for this project. Only a small portion of the integrative tools necessary for the effective application of omics techniques in clinical practice. (Abidalkareem et al., 2024).

Lipidomics:

A novel and promising technique for predicting and identifying breast cancer in high-risk patients is lipidomics. Lipidomics, a promising detection method, may be employed as a novel breast cancer prediction approach in high-risk patients. Since its inception in 2003, lipidomics has advanced significantly due to the development of mass spectrometry (MS) (Blanksby et al., 2010; Han et al., 2012). The field of lipidomics has been the subject of numerous recent publications, each with a distinct emphasis. Timely review of these factors is essential due to the quick development of technology and its applications. First, we go through MS-based methods for lipid analysis. Prior to mass spectrometric analysis, these methods differ depending on whether or not liquid chromatography (LC) is used. They may also differ in their analytical coverage, such as a "targeted" vs a "global" analysis, for instance in breast cancer (Han et al., 2012).

The purpose of this work was to use electrospray ionization tandem mass spectrometry in conjunction with triple quadrupole liquid chromatography to describe the global lipid profiles of 194 plasma samples. This study used electrospray ionization tandem mass spectrometry and triple quadrupole liquid chromatography to characterize the global lipid profiles of 194 plasma samples. These samples came from 110 people developed benign breast disease and 84 patients diagnosed with early-stage breast cancer (stages 0–II). To allow for complete analysis, the samples were separated into two sets: training and validation. The objective of this study was to determine the global lipid profiles of 194 plasma samples from 110 patients with benign breast disease and 84 patients with early-stage breast cancer (stages 0–II). These samples were divided into training and validation sets using triple quadrupole liquid chromatography electrospray ionization tandem mass spectrometry. To enable for more thorough evaluation, the samples were divided into two sets: training and validation. Using a mixture of 15 lipid types, the prediction model demonstrated significant diagnostic utility. In the training set, this combination yielded 83.3% sensitivity, 92.7% specificity, 89.7% positive predictive value (PPV), and 87.9% negative predictive value (NPV). The area under the curve (AUC) for the training

group was determined to be 0.926, with a 95% confidence interval of 0.869 to 0.982 (Blanksby et al., 2010; Han et al., 2012).

Omics for personalised medicine in breast cancer:

A thorough analysis of laboratory data, including tumor markers and immunohistochemical results as well as DNA and RNA data from whole-exome and whole-transcriptomic sequencing was carried out by the multi-omics clinical tumor board. A suggested treatment plan was developed after this research, which can involve taking part in clinical trials, using off-label medications, or getting extra care **(**Kato et al., 2020)**.**With a median age of 56 and a range of ages from 32 to 75, the study population was primarily made up of female volunteers. Eighty-nine percent of the participants, or a sizable majority, identified as White. Additionally, these patients had undergone a range of zero to thirteen lines of treatment, with a median of three. The multi-omics clinical tumor board received laboratory data like immunohistochemistry and tumor markers as well as DNA- and RNA-level data, like whole-exome sequencing and whole-transcriptomic sequencing **(**Massard et al., 2017). Optimal treatment management was then advised, which included clinical trials, off-label therapy, or additive care. Patients enrolled in the study were predominantly female, had a median age of 56 years (ranging from 32 to 75), a majority of white population (89%), and a median of 3 previous lines of treatment (ranging from 0 to 13). The liver (27%), lymph nodes (35%), bone (10%), and soft tissue (11%) accounted for the majority of biopsies (n = 63). Fifteen people in all had serial biopsies; each participant had two or more samples taken. The majority of biopsies (n = 63) came from the liver (27%), lymph nodes (35%), bone (10%), or soft tissue (11%). Fifteen individuals underwent serial biopsies, in which two or more samples were taken **(**Kato et al., 2020**;** Massard et al., 2017**).** Twelve people in all got matched therapy, yielding a PFS2:PFS1 ratio of 1.3. According to this ratio, the first progression-free survival (PFS1) was 1.3 times lower than the second progression-free survival (PFS2). Interestingly, eight of these patients had PFS2:PFS1 ratios that were either within or beyond the predetermined cutoff. Kong et al. provided a thorough analysis of the results for three patients who received matched therapy in their trial. In the first instance, the patient was treated with tamoxifen when the disease recurred after being first diagnosed as estrogen receptor (ER)-positive (Kong et al., 2021). In this case, tumor heterogeneity was detected by serial biopsies. Later liver biopsies with mixed results showed few to have triple-negative and ER-negative histological subtypes. Regrettably, the treating physicians could not identify a course of treatment that worked for this individual. In another case, a patient undergoing sequential biopsy monitoring was given a unique combination of a checkpoint and a PARP inhibitor. A significant development was the identification of resistance mechanisms within the MAP kinase pathway, which prompted the decision to switch the patient to bridging therapy with chemotherapy for tumor control. The patient was treated with a PARP inhibitor along with the MEK inhibitor cobimetinib (Cotellic) after a comprehensive debate at the tumor board **(**Mertins et al., 2016). Furthermore, while on doxorubicin therapy, another patient showed a remarkable tumor remission at first, but later developed a growing hepatic lesion and increased tumor markers. It was shown that this lesion possessed an activating mutation in ERBB3 and was both HER2-negative and ER-negative. Additional evidence of activation along the HER2 signaling axis came from research analytics. Consequently, the patient's tumor markers significantly decreased once trastuzumab (Herceptin) was introduced **(**Kong et al., 2021).

Matched Therapy Outcomes

Case	Key Mutation / Finding	Therapy	Outcome
Case 1	ER heterogeneity	Tamoxifen	No benefit
Case 2	MAPK resistance	PARP + MEK-inhibitor	Partial benefit
Case 3	ERBB3 mutation	Trastuzumab	High marker reduction

Table 3:- This table shows results of matched therapy a, emphasizing the treatment responses seen in individuals who received treatments customized based on certain molecular or biomarker profiles.

Current challenges

Despite its incredible promise, BC research on AI still faces several hurdles in order to bring its use in clinical practice to fruition. The main obstacle is the constraint of model validation. Many modern studies validate imaging genomics models using retrospective datasets from a single center. The limitations of existing models significantly hinder their ability to generalize and establish trust across various healthcare organizations. Furthermore, because these models have not been validated through prospective clinical trials, it is challenging to evaluate how well they perform in actual clinical settings and to compare their efficacy with the gold standard treatments currently in use **(**Hagiwara A et al., 2023)**.** It significantly reduces the models' dependability and capacity to generalize across various healthcare organizations. More significantly, it is difficult to assess the performance of current models in real clinical settings and compare their effectiveness to the current gold standard treatments because they have not been validated in prospective clinical studies. To assess AI models' clinical usefulness in practical applications, large multicenter, prospective trials are urgently needed more significantly, it is difficult to assess the performance of current models in real clinical settings and compare their effectiveness to the

current gold standard treatments because they have not been validated in prospective clinical studies. Significant obstacles to the models' therapeutic application in AI arise from their restricted interpretability **(**Cheng et al., 2025; Hagiwara et al., 2023)**.** Despite the excellent prediction accuracy of deep learning imaging genomics models, particularly the end-to-end models, their "black box" character makes it difficult for doctors to comprehend the decision-making process. This lack of openness may make people less trusting of them and less inclined to use them. Uncertainty surrounds the connection between tumor biological events and manually derived characteristics. Even in conventional imaging genomics models, the relationship between biological processes in tumors and manual features remains unclear. This uncertainty limits the model's ability to explain biological phenomena. A crucial area for further investigation is how to make AI models more interpretable and create models that are more transparent and simpler to comprehend. Last but not least, system integration and multimodal data integration are also problematic (Demircioglu et al., 2023; Wei et al., 2019).The integration of multidimensional data, including genomes, imaging genomics, and pathomics, is a major trend in the development of complete radio-multi-omics models. However, this endeavor necessitates the resolution of various technological challenges related to the integration and analysis of data sourced from diverse origins **(**Castiglioni et al., 2021)**.** A possible future trend for creating thorough radio-multi-omics models is the integration of multidimensional data, including as genomics, imaging genomics, and pathomics. However, in order to properly evaluate data from many sources, this integration poses technological problems that must be resolved. Furthermore, multi-omics data may be combined by artificial intelligence to improve patient stratification and predict treatment resistance (Castiglioni et al., 2021; Demircioglu et al., 2023).

Future directions

The following are the primary areas where future AI development in British Columbia should focus to overcome the aforementioned obstacles. Validation methods are the first step to prioritizing optimization. For this reason to increase the models' clinical usefulness and generalizability, multicenter validation should be given top priority in future research, and large, multicenter cohorts or publically available databases should be actively used for external validation **(**Hagiwara A et al., 2023**).**To generate more solid evidence for clinical application, it is crucial to actively support prospective clinical trials in the interim. Models that have done well in retrospective investigations can be elevated to a higher degree of evidence. The therapeutic utility of these models ought to be carefully investigated and contrasted with existing treatment approaches (Blanksby et al., 2010; Han et al., 2012)**.** Second, in order to gain the confidence of clinicians, it is essential to increase the model's interpretability. Secondly, improving the interpretability of the model is crucial to winning over physicians. The interpretability of radiomics based on deep learning (DL) should be studied in the future. This could entail pulling features from DL and combining them with conventional machine learning (ML) techniques, or it could entail employing technical techniques to examine how DL models make decisions (Hagiwara et al., 2023). By doing this, we can maintain the models' predicted accuracy while improving their interpretability. Examining the inherent relationships between imaging characteristics and tumor biological activity is part of this approach, which also entails combining different omics data. The objective is to increase the biological explanatory power of the model, which can lead to better treatment approaches. We hope to increase the model's capacity to offer biological insights and ultimately enhance treatment results by examining the connections between imaging features and tumor biology. To effectively use AI in healthcare settings, an integrated AI-assisted system must be created. To advance the function of AI in various situations, such a system must be developed (Topol et al., 2019). Future studies should concentrate on developing a multidisciplinary strategy that integrates different functional modules. The practical uses of Computer-Aided Diagnosis (CAD) systems in breast cancer screening should serve as inspiration. This includes tumor categorization, prognosis prediction, and treatment response evaluation. It is crucial to focus on making the system user-friendly by simplifying the interface for both doctors and patients. User-friendly design, comprehensive and understandable results, and, ultimately, the simple incorporation of AI technology into BC diagnosis and treatment procedures can all help achieve this target (Cheng J et al., 2025; Topol et al., 2019). Eventually, the system will be able to seamlessly integrate AI technology into the BC diagnostic and treatment process, dramatically improving the quality of BC diagnosis and therapy while increasing patient survival. The system will eventually be able to smoothly integrate AI technologies into the diagnosis and treatment of BC, which will significantly raise the quality of BC diagnosis and treatment as well as enhance patient survival (Cheng J et al., 2025; Hagiwara et al., 2023).

Conclusion

The integration of artificial intelligence into healthcare, with a focus on multi-omics data analysis, was the main subject of this essay. Healthcare systems can improve patient outcomes by properly integrating several omics while addressing ethical considerations like as privacy, openness, and fairness. In this study, how artificial intelligence (AI) is developing to manage the complexity of biological data and assessed existing deep learning techniques has been highlighted. Healthcare organizations may fully utilize multi-omics integration to enhance patient care by striking a balance between cutting-edge AI techniques and ethical issues. The state-of-the-art deep learning techniques has been examined and demonstrated how AI is developing to satisfy the needs of biological data. Non-generative models that help with

categorization and prediction include feedforward neural networks and graph attention networks. Variational autoencoders and other generative models, on the other hand, make data augmentation and synthesis easier. By combining, these methods, an increase is noticed in the breadth and precision of healthcare insights as long as data integrity and ethical norms are respected. The technological approaches used to handle each omics modality has been compared using data mainly from the Cancer Genome Atlas (TCGA). Despite the excellent precision of multi-omics integration shown by all the models examined, some are more appropriate for datasets that are incomplete. Specifically, VAEs and GCNs offer significant advantages in addressing the challenge of missing modalities. Nevertheless, the fairness indicators, subgroup-specific assessments, and external validity of many models are lacking, which restricts their applicability to different groups. Federated transformers and hybrid architectures are examples of new paradigms that are enhancing the resilience of multi-omics models. The resilience of multi-omics models is being improved by new paradigms such as federated transformers and hybrid architecture. Frameworks like NVIDIA FLARE provide privacy-preserving training across many universities, which is crucial for researching uncommon malignancies. Transformers use a pre-trained biological basis and cross-modal attention to help overcome integration problems. PATH-GPTOMIC better handle integration problems, transformers make use of this pretrained biological foundation and cross-modal attention. The integration of molecular and route data is the goal of recent developments in GPT-based survival modelling, where key issues in translational research are directly addressed by these advancements. Because transformer-based methods may capture long-range dependencies and replicate cross-modal interactions, researchers are investigating them. Federated learning is also becoming more well-known as a decentralized, privacy-conscious approach that works well for multi-institutional research with limited access to data. The study of transformer-based methods emphasizes how they may reproduce long-range interdependence and cross-modal interactions. These developments highlight the need for models that are safe, scalable, accurate, and prepared for practical use. Federated learning is becoming more widely acknowledged as a decentralized and privacy-conscious method, particularly advantageous for multi-institutional research with restricted data access. The necessity for models that are safe, scalable, accurate, and prepared for real-world applications is highlighted by these developments. In conclusion, multi-omics analysis relies heavily on effective pre-processing. It enables each data modality to be adapted to its distinct statistical properties, promoting strong integration and allowing customisation depending on each modality's individual qualities. Latent embedding frameworks, enhancing robustness and interpretability and recent advancements in models incorporate modality-specific pre-processing to realize the full potential of AI in precision medicine, it will be necessary in the future to integrate various types of data and create hybrid models**.** To effectively integrate computational techniques with medical practice, future research should concentrate on performing ablation trials, enhancing interpretability, and guaranteeing clinical validation across a range of patient demographics.

Author's contribution: SS conceived the idea and designed the framework. SS, drafted the manuscript. The author approved the final version of the manuscript.
Funding N/A.
Availability of data and material N/A.
Code availability N/A.

Declarations
Conflict of interest None.
Ethical approval N/A.
Consent to participate N/A.
Consent for publication N/A.

References

1. Abidalkareem A, Ibrahim AK, Abd M, Rehman O, Zhuang H. Identification of Gene Expression in Different Stages of Breast Cancer with Machine Learning. *Cancers (Basel)*. 2024;16(10):1864. Published 2024 May 14. doi:10.3390/cancers16101864
2. Baciu C, Xu C, Alim M, Prayitno K, Bhat M. Artificial intelligence applied to omics data in liver diseases: Enhancing clinical predictions. *Front Artif Intell*. 2022;5:1050439. Published 2022 Nov 15. doi:10.3389/frai.2022.1050439
3. Blanksby SJ, Mitchell TW. Advances in mass spectrometry for lipidomics. *Annu Rev Anal Chem (Palo Alto Calif)*. 2010;3:433-465. doi:10.1146/annurev.anchem.111808.073705
4. Blixt L, Bogdanovic G, Buggert M, et al. Covid-19 in patients with chronic lymphocytic leukemia: clinical outcome and B- and T-cell immunity during 13 months in consecutive patients. *Leukemia*. 2022;36(2):476-481. doi:10.1038/s41375-021-01424-w

5. Budczies J, Brockmöller SF, Müller BM, et al. Comparative metabolomics of estrogen receptor positive and estrogen receptor negative breast cancer: alterations in glutamine and beta-alanine metabolism. *J Proteomics*. 2013;94:279-288. doi:10.1016/j.jprot.2013.10.002
6. Castiglioni I, Rundo L, Codari M, et al. AI applications to medical images: From machine learning to deep learning. *Phys Med*. 2021;83:9-24. doi:10.1016/j.ejmp.2021.02.006
7. Castiglioni I, Ippolito D, Interlenghi M, et al. Machine learning applied on chest x-ray can aid in the diagnosis of COVID-19: a first experience from Lombardy, Italy. *Eur Radiol Exp*. 2021;5(1):7. Published 2021 Feb 2. doi:10.1186/s41747-020-00203-z
8. Chaudhary A, Gustafson D, Mathys A. Multi-indicator sustainability assessment of global food systems. *Nat Commun*. 2018;9(1):848. Published 2018 Feb 27. doi:10.1038/s41467-018-03308-7
9. Cheng, L., & Yeap, Q. F. (2025). The 2025 Youth Transition Report: Outcomes for youth and young adults with disabilities. Institute for Educational Leadership. https://iel.org/iel-youth-transition-reports-outcomes-for-youth-and-young-adults-with-disabilities-2025.
10. Ching T, Zhu X, Garmire LX. Cox-nnet: An artificial neural network method for prognosis prediction of high-throughput omics data. *PLoS Comput Biol*. 2018;14(4):e1006076. Published 2018 Apr 10. doi:10.1371/journal.pcbi.1006076
11. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559-1567. doi:10.1038/s41591-018-0177-5
12. Demircioglu T, Karakus M, Ucar S. Developing Students' Critical Thinking Skills and Argumentation Abilities Through Augmented Reality-Based Argumentation Activities in Science Classes. *Sci Educ (Dordr)*. Published online August 22, 2022. doi:10.1007/s11191-022-00369-5
13. Duan GL, Zhou Y, Tong YP, Mukhopadhyay R, Rosen BP, Zhu YG. A CDC25 homologue from rice functions as an arsenate reductase. *New Phytol*. 2007;174(2):311-321. doi:10.1111/j.1469-8137.2007.02009.x
14. Geyer PE, Arend FM, Doll S, et al. High-resolution serum proteome trajectories in COVID-19 reveal patient-specific seroconversion. *EMBO Mol Med*. 2021;13(8):e14167. doi:10.15252/emmm.202114167
15. Hagiwara S, Shiohama T, Takahashi S, et al. Comprehensive High-Depth Proteomic Analysis of Plasma Extracellular Vesicles Containing Preparations in Rett Syndrome. *Biomedicines*. 2024;12(10):2172. Published 2024 Sep 24. doi:10.3390/biomedicines12102172
16. Han, J., Kamber, M., & Pei, J. (2012). Data mining: Concepts and techniques (3rd ed.). Waltham, MA: Morgan Kaufmann Publishers. DOI: 10.1016/C2009-0-61819-5
17. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol*. 2017;18(1):83. Published 2017 May 5. doi:10.1186/s13059-017-1215-1
I. H., et al. (2021). Artificial intelligence in cancer diagnosis and therapy: A comprehensive review. *Journal of AI in Medicine, 12*(4), 102–115.
18. Disis MLN, Guthrie KA, Liu Y, et al. Safety and Outcomes of a Plasmid DNA Vaccine Encoding the ERBB2 Intracellular Domain in Patients With Advanced-Stage ERBB2-Positive Breast Cancer: A Phase 1 Nonrandomized Clinical Trial. *JAMA Oncol*. 2023;9(1):71-78. doi:10.1001/jamaoncol.2022.5143
19. Jiang Y, Sun C, Xu J, et al. Synthesis-on-substrate of quantum dot solids. *Nature*. 2022;612(7941):679-684. doi:10.1038/s41586-022-05486-3
20. Jinthanasatian, P., Auephanwiriyakul, S., & Theera-Umpon, N. (2017). Microarray data classification using neuro-fuzzy classifier with firefly algorithm. In 2017 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 1–6). Honolulu, HI, USA: IEEE. DOI: 10.1109/SSCI.2017.8285324
21. Johansson, R. (2019). Numerical Python: Scientific computing and data science applications with NumPy, SciPy and Matplotlib (2nd ed.). Apress. https://doi.org/10.1007/978-1-4842-4246-9
22. Jung KW, Kang MJ, Park EH, et al. Prediction of Cancer Incidence and Mortality in Korea, 2023. *Cancer Res Treat*. 2023;55(2):400-407. doi:10.4143/crt.2023.448
23. Kather JN, Heij LR, Grabsch HI, et al. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer*. 2020;1(8):789-799. doi:10.1038/s43018-020-0087-6
24. Kato TA, Sartorius N, Shinfuku N. Forced social isolation due to COVID-19 and consequent mental health problems: Lessons from hikikomori. *Psychiatry Clin Neurosci*. 2020;74(9):506-507. doi:10.1111/pcn.13112
25. Kim S, Chen J, Cheng T, et al. PubChem 2023 update. *Nucleic Acids Res*. 2023;51(D1):D1373-D1380. doi:10.1093/nar/gkac956
26. Kong J, Shin Y, Röhr JA, et al. CO2 doping of organic interlayers for perovskite solar cells. *Nature*. 2021;594(7861):51-56. doi:10.1038/s41586-021-03518-y
27. Kristensen VN, Lingjærde OC, Russnes HG, Vollan HK, Frigessi A, Børresen-Dale AL. Principles and methods of integrative genomic analyses in cancer. *Nat Rev Cancer*. 2014;14(5):299-313. doi:10.1038/nrc3721

28. Li YH, Yu CY, Li XX, et al. Therapeutic target database update 2018: enriched resource for facilitating bench-to-clinic research of targeted therapeutics. *Nucleic Acids Res*. 2018;46(D1):D1121-D1127. doi:10.1093/nar/gkx1076
29. Li Z, Li S, Luo M, et al. dbPTM in 2022: an updated database for exploring regulatory networks and functional associations of protein post-translational modifications. *Nucleic Acids Res*. 2022;50(D1):D471-D479. doi:10.1093/nar/gkab1017
30. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng*. 2021;5(6):555-570. doi:10.1038/s41551-020-00682-w
31. Massard C, Michiels S, Ferté C, et al. High-Throughput Genomics and Clinical Outcome in Hard-to-Treat Advanced Cancers: Results of the MOSCATO 01 Trial. *Cancer Discov*. 2017;7(6):586-595. doi:10.1158/2159-8290.CD-16-1396
32. Mertins P, Mani DR, Ruggles KV, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature*.2016;534(7605):55-62. doi:10.1038/nature18003
33. Rahaman, M. M., Islam, M. R., Bhuiyan, M. M. R., Noman, I. R., Aziz, M. M., & Das, K. (2025, March). Harnessing big data in biotechnology: A machine learning approach to multi-omics. In *2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)* (pp. 1391-1401). IEEE.
34. Subramanian I, Farahnik J, Mischley LK. Synergy of pandemics-social isolation is associated with worsened Parkinson severity and quality of life. *NPJ Parkinsons Dis*. 2020;6:28. Published 2020 Oct 8. doi:10.1038/s41531-020-00128-9
35. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7