International Journal of Research in Medical Sciences
Manoj M et al. Int J Res Med Sci. 2024 Feb;12(2):641-647
www.msjonline.org
pISSN 2320-6071 | eISSN 2320-6012

DOI: https://dx.doi.org/10.18203/2320-6012.ijrms20240244

Review Article

Advancements in early cancer diagnosis using blood-based biomarkers and a machine learning approach

Megha Manoj, Tharunya Kalva Banuprakash, Namasivaya Naveen Shanmuga Sundaram*, Rajive Gandhi Chandrabose, Dharshini Sankareshwaran, Rakshaana Behum Kadar Mohideen

Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research, Porur, Chennai, Tamil Nadu, India

Received: 04 December 2023
Revised: 04 January 2024
Accepted: 15 January 2024

***Correspondence:**
Namasivaya Naveen Shanmuga Sundaram,
E-mail: namasivayanaveen@gmail.com

ABSTRACT

Mutations that promote aberrant cell growth are the root of the condition known as cancer. There are over a hundred distinct forms of cancer that have been identified, with lung, colon, pancreatic, breast, kidney, and prostate cancer being the most prevalent. The likelihood that a patient will survive cancer is significantly improved by early identification. Most techniques used to detect cancer are invasive, which may be painful and uncomfortable for patients and prevent them from seeking treatment. As a result, cancer is frequently discovered only after substantial symptoms have developed and it may then be too late for treatment. In this review, we will discuss several methods for detecting cancer through blood tests, different elements that serve as biomarkers, and machine learning algorithms for predicting outcomes.

Keywords: Early cancer detection, Blood, ctDNA, Machine learning

INTRODUCTION

Cancer is a term used to denote a collection of diseases caused due to mutation, leading to the abnormal proliferation of cells.[1] There are greater than hundred unique types of documented cancer among them the most common cancers include lung, colorectal, breast, kidney, prostate, and bladder cancer.[1,2] The causative agent of cancer is known as a carcinogen and it includes oncogenic viruses, radiation, and chemicals.[3] Appropriate detection methods are necessary for effective cancer detection and treatment. Imaging tests produce detailed images of regions inside the body and the different energy forms such as radio waves, x-rays, radioactive substances, and ultrasound are commonly used in imaging tests. They are mainly used in the detection and treatment of diseases, and to monitor the treatment progress. Biopsy is the process in which a part of a tissue or cells or liquid is extracted from the body and is examined for the presence of a tumor. Biopsy is the most common and predominantly used method to detect the type of cancer. The endoscopy procedure uses a thin, flexible tube containing a camera called an endoscope which is inserted into the body to detect abnormal areas. If any abnormality is found, biopsy is done and the presence of the tumor is detected.

One of the leading causes of death is cancer which accounts for 10 million deaths worldwide. We can overcome this situation if the cancer is detected at an early stage. Early detection has a great impact on increasing the patient's survival rate. The morbidity rate can be minimized when the cancer is detected early. Cancer treatment has a high success rate when detected early which greatly increases the survival rate.[4]

In many cases, cancer is detected only after significant symptoms have developed when it may be too late for treatment and most of the methods available for detection of cancer are invasive which may be uncomfortable and painful and can discourage patients.

The DNA released by tumor cells into the blood is known as circulating tumor DNA or ctDNA the old cells are replaced by new cells as the tumor develops the dead cells are broken-down and their components are given out into the blood in this way the tumor DNA enters the blood and spreads all over the body thence spreading the tumor. The ctDNA acts as a biomarker and helps in the detection of tumors. Once the released ctDNA is isolated from other components, machine learning algorithms can be applied and the presence of tumor can be detected. The tumor stage can be identified by the amount and type of ctDNA released into the bloodstream.[5]

BEAMing stands for Bead Emulsion Amplification and magnetics. TAm-Seq is a sequencing method that is used to identify rare mutations in cfDNA. CAPP-Seq is used in the quantification of mutations in ctDNA. WGS and WES are deep sequencing of entire genome and exome respectively. The Table 1 represents the different methods that are widely used to detect ctDNA. ddPCR stands for digital polymerase chain reaction and it uses absolute quantification method to detect the ctDNA. WGBS-Seq is used in the detection of DNA methylation patterns.[6] Hence, appropriate methods to detect cancer are of utmost importance. Recent studies have shown that with the help of machine learning, it is possible to detect and diagnose over 50 types of cancer with a simple blood test at the earliest stage before any symptoms develop.[7] In this review, we focus mainly on the several methods for cancer detection from blood test, various substances that, and machine learning algorithms for predicting outcomes.

Table 1: Some of the ctDNA detection methods.

Methods	Description
ddPCR	Absolute quantification of the sample
BEAMing	Bead, emulsion, amplification, and magnetics
TAm-Seq	It is a sequencing method to detect rare mutations in cfDNA
CAPP-Seq	It is a technique used in the detection and quantification of mutations in ctDNA
WGS	Deep sequencing of the entire genome
WES	Deep sequencing of the exome
WGBS-Seq	It is used to detect DNA methylation patterns

LIQUID BIOPSY COMPONENTS

Liquid biopsy (LB) is a minimally-invasive platform for cancer diagnosis and molecular monitoring achieved by a powerful real-time approach. Coined by Pantel and Alix-Panabieres, it is used in biomarker detection in blood and is also used in the study of Circulating Tumor Cells (CTCs) and other components. LB components can be extracted from blood and other body fluids like urine, saliva, CSF, etc.[8] Blood is the most commonly used biofluid in liquid biopsy due to the presence of CTCs and nucleic acids which migrate from the primary tumor sites to other locations of the body.

The presence of tumor marker proteins like Prostate-specific antigen (PSA), Carbohydrate antigen 19-9 (CA19.9), and Carcinoembryonic antigen (CEA) and blood biomarkers like Circulating Tumor RNA (ctRNA) and Circulating Tumor DNA (ctDNA) with the help of highly sensitive genomic technologies make the detection process much easier. The BRCA1, BRCA2, TP53, and PALB2 are some important gene mutations detected in different tissues that cause cancer. The Table 2 explains the various gene mutations that occur in different cancer types.

Table 2: Different types of cancer and their respective gene mutations.

Type of cancer	Type of gene mutation
Breast cancer in women[9]	ATM, BARD1, BRCA1, BRCA2, BRIP1, CHEK2, CDH1, NF1, PALB2, PTEN, RAD51C, RAD51D, STK11, TP53
Breast cancer in men[10]	BRCA1, BRCA2, CHEK2, PALB2
Ovarian cancer[11]	ATM, BRCA1, BRCA2, BRIP1, EPCAM, MLH1, MSH2, MSH6, NBN**, PALB2, RAD51C, RAD51D
Oral cancer[12]	Tumor suppressor gene p53
Lung cancer[13]	TP53, EGFR, and KRAS genes
Colorectal cancer[14]	APC, EPCAM, MLH1, MSH2, MSH6, PMS2, CHEK2, PTEN, STK11, TP53, MUTYH
Pancreatic cancer[15]	ATM, BRCA1, BRCA2, CDKN2A, EPCAM**, MLH1, MSH2**, MSH6**, PALB2, STK11, TP53

Note: **BRCA1 is a type of mutations which have been linked to a very dangerous and very rare type of caner known as serous endometrial cancer, **-implies that more research is needed to confirm a link between this cancer type and an inherited mutation in this gene.

Mutations in the KRAS gene have been detected in patients with pancreatic and colorectal cancer during urine DNA analysis. Similarly, sensitive mutations in epidermal growth factor receptors (EFGR) have been detected in non-small-cell lung cancer patients. Lung cancer can be detected through saliva using mRNA as a biomarker.

Some proteins like calprotectin, zinc 2-glycoprotein (AZGP1), and haptoglobin hp2 (HP) were expressed in high levels in the saliva of lung cancer patients.[8] There are six major components of Liquid Biopsy which are as follows, Circulating Tumor Cells (CTCs), Cell-Free DNA (cfDNA), Circulating Tumor DNA (ctDNA), Cell-Free

RNA (cfRNA), Exosomes, and Tumor Educated Platelets (TEPs).[8-10]

CIRCULATING TUMOR CELLS

An Australian scientist Thomas Ashworth first observed CTCs in 1869. He found primary tumor-like cells in the bloodstream of an autopsied metastatic cancer patient.[9] Migrating cancer cells that release or shed actively into the blood from their primary location are called CTCs. Even though CTCs contents are rare, they are differentiated from other blood cells with the help of various positive and negative expressions of markers. Hence, they can be used as an alternative to invasive biopsy methods for cancer detection. CTC isolation from peripheral blood can avoid complex invasive biopsy processes. Studies show that CTC-derived xenografts reflect more cancer characteristics and genetic diversity than homogenous tumor cell culture. CTCs are extremely limited in number. Generally, isolated CTCs purity is less than 1%.[8]

Clinical laboratories use various methods to determine circulating tumor cells (CTCs). These methods include Oncoquick, which is based on density gradient, and ISET, which uses morphological techniques. Immunomagnetic techniques such as MACS®, AdnaTest®, RARETM®, FAST®, LSC®, CellSearch®, and CellCollector® are also commonly employed to detect CTCs. Moreover, automated systems such as microscopes like ACIS®, CellSpotter®, and ARIOL® can provide further exploration of CTCs. These different methods allow for efficient and accurate detection of CTCs, which can provide valuable information for cancer diagnosis, prognosis, and treatment. There are some limitations to CTCs as a biomarker. There is a need for a reasonable and effective enrichment method. Obtaining a sufficient number of CTCs for evaluation is one of the main challenges. Improvement in techniques used for assessing molecular characteristics of CTCs is needed.[9]

CELL-FREE DNA

Mandel and Metais first identified the cell-free DNA in blood plasma in the year 1948. Due to limitations in technologies, it took around 40 years to experimentally prove the role of cfDNA in cancer patients.[11,12] cfDNA denotes to the DNA that is actively given out into the bloodstream by apoptotic cells. Part of cfDNA in cancer victims comes from tumor cells.[9,13]

Not every cfDNA identified in the blood is tumor-derived; some may have originated in blood cells and others may have originated in other tissues. In non-cancer persons, they generally arise from apoptosis of cells. Patients with cancerous tumors have greater cfDNA levels or concentrations than that of patients with noncancerous tumors. Different methods have been developed for extracting cfDNA from the blood plasma. (i) magnetic beads and (ii) silica-based are two methods commonly used in the extraction of cfDNA. Liquid biopsy techniques are applied to detect specific mutations that act as an essential marker for targeted cancer.[8]

cfDNA EXTRACTION METHOD

Magnetic beads method of cfDNA extraction is one of the most reliable methods to extract cfDNA. In this method, positive magnetic beads bind to negative charged phosphate backbone of DNA, while other particles remain in solution. The main benefit of this approach is that no centrifugation or vacuum manifolds are required, which can be a challenging task in many automated processes. Separation can be performed manually, or it can be semi-automated, or fully automated. Once the DNA is extracted, it can be sequenced and analyzed.[14] In Silica-based method, the DNA binds to the silica under high-salt conditions. Chaotropic salts are the major element in extracting cfDNA. Nucleic acid binding to silica is facilitated by high concentrations of chaotropic salts. The nucleic acid is washed after the DNA has been bound to the silica membrane which washes away unwanted particles from the solution. Finally, the DNA is eluted under low salt concentration. Silica-based method are simple to perform, fast and economical. The Table 3 contains details of cfDNA extraction kits based on magnetic beads and silica-based methods. QIAamp Circulating Nucleic Acid Kit (CNA) and Zymo quick ccfDNA serum and plasma kit (Z) uses silica-based method and takes 1-4 ml and <10 ml of solution of input and can elute around 20-150 µl and >50 µl of solution respectively. Maxwell RSC ccfDNA Plasma Kit (RSC) and QIAamp MinElute ccfDNA midi kit (ME) are based on magnetic beads method and can elute 50 µl and 20-80 µl of solution respectively.[14]

Table 1: Some widely used cfDNA extraction kits.

Kit	Method	Input (ml)	Elution (µl)	Exe-cution
QIAamp circulating nucleic acid kit (CNA)	Silica-based	1-4	20-150	Manual
Maxwell RSC ccfDNA plasma kit (RSC)	Magn-etic beads	0.2-1	50	Automated
Zymo quick ccfDNA serum and plasma kit (Z)	Silica-based	<10	>50	Manual
QIAamp Min elute ccfDNA midi kit (ME)	Magn-etic beads	4-10	20-80	Manual

IMPROVING CFDNA EXTRACTION PURITY

Liquid biopsy is a minimally invasive or non-invasive method of detecting circulating cell-free DNA (cfDNA), which can be used to detect cancer at an early stage. The presence of residual impurities greatly influences DNA yield and quality. As a result, prior to analysis, the cfDNA must be purified. A PEG-based DNA clean-up can effectively remove impurities from extracted cfDNA. An aqueous nucleic acid sample is incubated with a solution containing salt and a specific concentration of PEG (PEG8000). This causes the nucleic acids to precipitate, after which they are pelleted by centrifugation and washed twice in 70% ethanol and pure cfDNA is collected. The purity of cfDNA can be significantly increased in this manner.[15]

CIRCULATING-TUMOR DNA

Tumor DNA that circulates in plasma is denoted as circulating tumor DNA or ctDNA, it can be a single-stranded or a double-stranded DNA. Early research revealed that ctDNA showed numerous cancer-related characteristics related to molecular changes, such as SNPs, methylation alterations, and thus was predicted to be originated from cancer tissue. There are three major sources for ctDNA: CTCs, live cancer cells, and apoptotic or necrotic tumor cells.

The rate of circulating-tumor DNA or ctDNA released into the bloodstream of the body is determined by the tumor's location, size, and vascularity, resulting in different levels of ctDNA ranging from blood from 0.01 percent to 90%.[8] In cancer patients the concentration of cfDNA can reach 1000 ng/ml.[8] ctDNA is present in low concentrations (0.01% of the total cfDNA) in the blood, making the detection challenging at the early stages of cancer.[11,16] Molecular-biology techniques like BEAMing, Q-PCR, CAPP-Seq, Safe-SeqS, and TAmSeq are sensitive, specific, and cost-efficient detection methods which uses few tumor-specific mutations.[8,17]

CELL-FREE RNA

Intracellular non-coding RNA molecules (which contain about 22 nucleotides) called miRNA (microRNA) were first discovered by Lee in 1993. They are important mediating signals in post-transcriptional silencing. Large amounts of cfRNA are generated due to the high turnover rate of tumor gene expression.[9,11]

As the concentration of miRNAs is higher than the mRNA and is more stable in cancer patients, it is identified easily in blood. Inconsistency in reference gene selection (internal or external) for quantitative detection is a limitation of using cfRNA as a biomarker. During the separation process, miRNAs show differences in quantity and quality due to extraction from different sources, which leads to unreliable results. Hence, isolation and quantification of miRNA and data analytic methods need more improvement.[8]

Exosomes

Exosomes are extracellular vesicles released from cells upon fusion of an intermediate endocytic compartment, the multivesicular body, with the plasma membrane. They contain constituents of the cells that secrete them such as protein, lipids, enzymes, DNA, and RNA.[18]

Tumor educated platelets

The second most abundant cells in peripheral blood platelets are circulating nucleated cell fragments that are about 2-3 µ in diameter. Studies show that platelets play a central role in tumor growth responses. Platelets interact with the cancer cells through a chemical process called sequestration of biomolecules associated with tumors, resulting in their 'education'. It affects the gene expression and alters the RNA profile of the platelets. Ongoing research suggests the use of tumor-educated platelets acts as an informative tool in the clinical diagnosis of cancer and also as an LB biomarker.[8,9]

COMPARISON OF THE ACCURACY OF DIFFERENT METHODS FOR DETECTING CANCER FROM BLOOD TEST

Different strategies and methods are under evaluation for the early detection of cancer using machine learning protocols. We explored different supervised learning approaches for multiple cancer type detection and compared their accuracies. The methods include GRAIL, CancerSEEK, PCR-based cobas EGFR Mutation test v2, Guardant360 CDX assay, PanSeer test (Singlera Genomics), and FoundationOne liquid CDX. It can be observed that FoundationOne Liquid CDX has a high accuracy of 95% in detecting cancer. The other methods such as Guardant360 CDX assay, PanSeer test (Singlera Genomics) have also reasonably good accuracies of greater than 80%. CancerSEEK and PCR-based cobas EGFR mutation test have an average accuracy between 60%-80% and GRAIL with an approximate of around 50%. The accuracy of each method varies from each other because of the different protocols and biomarkers they use for the detection.

The Figure 1 shows the accuracy comparison between different methods for cancer detection. Different colors represent different methods. Horizontal axis is ordered by different methods. The vertical axis is ordered by the accuracy value in percentage. Each bar represents the accuracy of each method on detecting various cancer types.[19] FoundationOne Liquid CDX which is based on next generation sequencing uses cfDNA with a copy number alteration in BRCA1, BRCA2, ERBB2 and rearrangements in ALK, BRCA1, and BRCA2 as biomarker and can detect multiple cancer types such as non-small cell lung cancer, Prostate cancer, Ovarian cancer, and Breast cancer. PanSeer test (Singlera genomics), non-invasive blood test based on DNA methylation patterns use ctmDNA (circulating tumor methylated DNA) as biomarker for detecting five common types of cancer (stomach, oesophageal, colorectal, ling and liver cancer).[20]

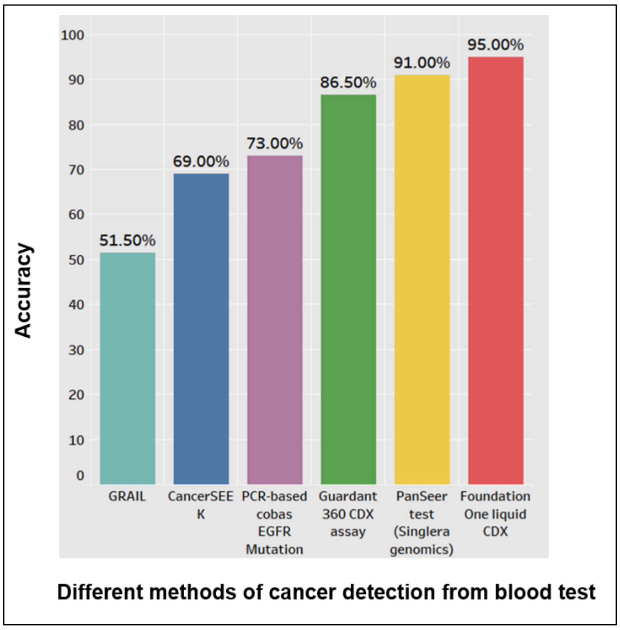

Figure 1: Comparison of the accuracies of different methods for detecting cancer from blood test.

Guardant360 CDX liquid biopsy detects KRAS p.G12C (oncogenic driver mutation in multiple cancer types) in patients with small-cell lung cancer.[21] Cobas EFGR Mutation Test v2 is a PCR based technique that detects mutations in exon 18 to 20 of EFGR (Epidermal Growth Factor Receptor) gene from small-cell lung cancer.[22] CancerSEEK which is based on random forest used ctDNA as biomarker and can detect five cancer types with a sensitivity >69% (ovarian, liver, gastric, pancreas, oesophageal cancers).[23] GRAIL detects mutations in cfDNA of blood plasma samples in 50 different cancer types with an average accuracy of 51.5% and it increases with the stage of cancer.[24]

MACHINE LEARNING IN CANCER DETECTION

Machine learning uses various probabilistic, statistical techniques which allow the machines to learn and gain information from past examples and experiences to identify and detect difficult, hidden, and complex patterns from large data. Hence, it is extremely applied in the detection and diagnosis of cancer.

There are different algorithms within machine learning and by comparing the accuracies the best-suited algorithm can be found.[25] Machine learning (ML) comprises of two main divisions as (i) supervised learning (ii) unsupervised learning. A set of labelled values is used as training data to evaluate the input. In contrast, no labelled examples are provided in unsupervised learning methods. It identifies and discovers patterns which are hidden or in groups available in the input. Apart from supervised and unsupervised learning, there is another type of method known as Semi-supervised learning, which combines and uses both supervised and unsupervised learning. It takes both labelled and unlabeled data into account to build an accurate learning model.[26]

ML techniques such as decision trees (DTs) and Artificial neural networks (ANNs) has been used in the detection of cancer for over three decades. Over the last decade, there has been an increase in the usage of other various supervised learning techniques, such as Support Vector Machines and Bayesian Networks, for cancer prediction and prognosis. The above-mentioned classification algorithms have been used cancer research widely.[26] ANNs are used to solve a wide range of classification and pattern recognition problems. This model combines the input variables to generate an output. Multiple unrevealed or hidden layers that mathematically constitute neural connections are generally used in ANNs. Although AANNs acts as a standard and appropriate method for classification problems, it has its own setbacks. Their layered architecture is time-consuming and can result in very poor performance.[27] Decision trees use a tree-structured classification pattern, with nodes representing input value and leaves representing decision output. Decision tree is one of the first and most well-known machine learning technique that has been widely used for classification. The architecture of DTs is designed in such a way that they are easy to understand and interpret. Their very-specific architecture-driven decisions allow for appropriate reasoning, which makes them a captivating technique.[26]

Support vector machines (SVM) are one of the most recent approaches in machine learning which is used in prediction and prognosis of cancer. SVM technique performs the classification by creating a 2-dimentional or 3-dimensional hyperplane that divides the given data points or values into two unique classes. The derived classifier has a high degree of generalizability and can thus be used to reliably classify new samples.[28] Bayesian classifiers generate probabilistic estimates instead of prediction. They indicate knowledge along with variables of interest having probabilistic dependencies using a directed acyclic graph. Bayesian Networks (BNs) has been progressively used in a variety of classification tasks and also in reasoning and knowledge representation. Apart from these majorly used techniques, algorithms such as LR and multi linear regression, logistic regression are also used in cancer studies.[26]

MACHINE-LEARNING PROCEDURE

Cancer detection is viewed as a supervised problem in machine learning. The general procedure of supervised machine learning consisting of seven steps is shown in the Figure 2. The input data is collected in this step. The quality and quantity of data are very important as they have a direct impact on the accuracy of the model. The data can be collected from an existing database or it can be created on our own. Data preparation includes handling missing values, normalization, feature extraction and train-test split.[9] There are different types of algorithms for different problems. Generally, an algorithm is chosen based on the type of problem. Since cancer detection is a classification problem, a suitable classification algorithm can be used.

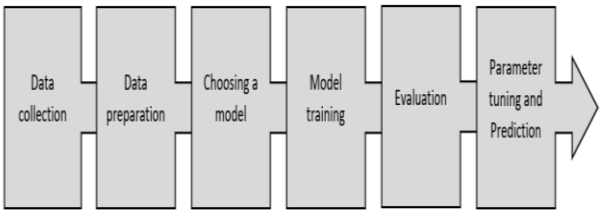

Figure 2: Workflow of a ML algorithm.[9]

Generally used classification methods or algorithms includes NBs, DTs, logistic regression (LR), K-nearest neighbors (KNNs), and SVMs. This is the stage at which the ML algorithm is trained using datasets fed into it. Appropriate and consistent training can boost the ML model's prediction rate. Once the model is trained, it is tested to see how well it performs in real-world scenarios. if the acquired results are not satisfactory, the previous steps are reanalyzed and the model's performance is improved. It is also known as hyperparameter tuning. This step focuses on improving the obtained results during evaluation. The final step in the machine learning process is prediction. In this step, the model is tested with unseen data and the efficiency is evaluated and the model is made ready to predict real-world scenarios.[9]

CONCLUSION

The use of machine learning in early cancer detection from blood tests holds immense potential for revolutionizing cancer diagnosis and treatment. With the help of advanced algorithms and techniques, it is possible to detect specific biomarkers in blood that indicate the presence of cancer cells at an early stage, even before the appearance of physical symptoms. This early detection can significantly improve patient outcomes and survival rates by enabling timely and accurate diagnosis and treatment. Moreover, the use of machine learning can also help reduce healthcare costs and improve resource allocation by identifying patients who require further diagnostic tests and treatment. As technology advances and more data becomes available, it is likely that machine learning-based methods will become even more effective in detecting cancer at an early stage, ultimately leading to better patient outcomes and a reduction in the burden of cancer worldwide.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: Not required

REFERENCES

1. Cooper GM, Hausman R. A molecular approach. The Cell. 2nd ed. Sunderland, MA: Sinauer Associates; 2000.
2. Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, et al. SEER Cancer Statistics Review, 1975-2010, National Cancer Institute. Available at: https://seer.cancer.gov/archive/csr/1975. Accessed on 22 November 2024.
3. Blackadar CB. Historical review of the causes of cancer. World J Clin Oncol. 2016;7(1):54-86.
4. Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. Lancet Glob Health. 2020;8(2):e180-90.
5. Cheng F, Su L, Qian C. Circulating tumor DNA: a promising biomarker in the liquid biopsy of cancer. Oncotarget. 2016;7(30):48832-41.
6. Li H, Jing C, Wu J, Ni J, Sha H, Xu X, et al. Circulating tumor DNA detection: A potential tool for colorectal cancer management. Oncol Lett. 2019;17(2):1409-16.
7. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV; CCGA Consortium. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol. 2020;31(6):745-59.
8. Fernández-Lázaro D, García Hernández JL, García AC, Córdova Martínez A, Mielgo-Ayuso J, Cruz-Hernández JJ. Liquid Biopsy as Novel Tool in Precision Medicine: Origins, Properties, Identification and Clinical Perspective of Cancer's Biomarkers. Diagnostics (Basel). 2020;10(4):215.
9. Liu L, Chen X, Petinrin OO, Zhang W, Rahaman S, Tang ZR, et al. Machine Learning Protocols in Early Cancer Detection Based on Liquid Biopsy: A Survey. Life (Basel). 2021;11(7):638.
10. Wang J, Chang S, Li G, Sun Y. Application of liquid biopsy in precision medicine: opportunities and challenges. Front Med. 2017;11(4):522-7.
11. Mandel P, Metais P. Nuclear Acids In Human Blood Plasma. C R Seances Soc Biol Fil. 1948;142(3-4):241-3.
12. Stroun M, Anker P, Maurice P, Lyautey J, Lederrey C, Beljanski M. Neoplastic characteristics of the DNA found in the plasma of cancer patients. Oncology. 1989;46(5):318-22.
13. Vendrell JA, Mau-Them FT, Béganton B, Godreuil S, Coopman P, Solassol J. Circulating Cell Free Tumor DNA Detection as a Routine Tool forLung Cancer Patient Management. Int J Mol Sci. 2017;18(2):264.
14. Lee H, Park C, Na W, Park KH, Shin S. Precision cell-free DNA extraction for liquid biopsy by integrated microfluidics. NPJ Precis Oncol. 2020;4:3.
15. Lin SY, Luo Y, Marshall MM, Johnson BJ, Park SR, Wang Z, et al. A New Method for Improving Extraction Efficiency and Purity of Urine and Plasma

Cell-Free DNA. Diagnostics (Basel). 2021;11(4):650.
16. Boyle P. Triple-negative breast cancer: epidemiological considerations and recommendations. Ann Oncol. 2012;23(6):vi7-12.
17. Hou J, Meng F, Chan LW, Cho WC, Wong SC. Circulating Plasma MicroRNAs As Diagnostic Markers for NSCLC. Front Genet. 2016;7:193.
18. Edgar JR. Q&A: What are exosomes, exactly? BMC Biol. 2016;14:46.
19. Wong KC, Chen J, Zhang J, Lin J, Yan S, Zhang S, et al. Early Cancer Detection from Multianalyte Blood Test Results. iScience. 2019 May 31;15:332-41.
20. Woodhouse R, Li M, Hughes J, Delfosse D, Skoletsky J, Ma P, et al. Clinical and analytical validation of FoundationOne Liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin. PLoS One. 2020;15(9):e0237802.
21. Bauml JM, Li BT, Velcheti V, Govindan R, Curioni-Fontecedro A, Dooms C, et al. Clinical validation of Guardant360 CDx as a blood-based companion diagnostic for sotorasib. Lung Cancer. 2022;166:270-78.
22. Han AL, Kim HR, Choi KH, Hwang KE, Zhu M, Huang Y, et al. Comparison of cobas EGFR Mutation Test v2 and PANAMutyper-R-EGFR for Detection and Semi-Quantification of Epidermal Growth Factor Receptor Mutations in Plasma and Pleural Effusion Supernatant. Ann Lab Med. 2019;39(5):478-87.
23. Killock D. Diagnosis: CancerSEEK and destroy - a blood test for early cancer detection. Nat Rev Clin Oncol. 2018;15(3):133.
24. Ofman JJ. GRAIL and the quest for earlier multi-cancer detection. Nat Res. 2020.
25. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. Cancer Inform. 2007;2:59-77.
26. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J. 2014;13:8-17.
27. Ayer T, Alagoz O, Chhatwal J, Shavlik JW, Kahn CE Jr, Burnside ES. Breast cancer risk estimation with artificial neural networks revisited: discrimination and calibration. Cancer. 2010;116(14):3310-21.
28. Platt JN, Cristianini N, Shawe-Taylor J. Large margin DAGs for multiclass classification. Adv Neural Inf Proces Syst. 1999;12.

Cite this article as: Manoj M, Banuprakash TK, Sundaram NNS, Chandrabose RG, Sankareshwaran D, Mohideen RBK. Advancements in early cancer diagnosis using blood-based biomarkers and a machine learning approach. Int J Res Med Sci 2024;12:641-7.