

Role of Machine Learning in Liquid Biopsy of Brain Tumours

Zanib Javed, Saqib Kamran Bakhshi, Saad Akhtar Khan, Muhammad Shahzad Shamim

Abstract

Liquid biopsy has multiple benefits and is used extensively in other fields of oncology, but its role in neuro-oncology has been limited so far. Multiple tumour-derived materials like circulating tumour cells (CTCs), tumour-educated platelets (TEPs), cell-free DNA (cfDNA), circulating tumour DNA (ctDNA), and miRNA are studied in CSF, blood (plasma, serum) or urine. Large and complex amounts of data from liquid biopsy can be simplified by machine learning using various algorithms. By using this technique, we can diagnose brain tumours and differentiate low versus high-grade glioma and true progression from pseudo-progression. The potential of liquid biopsy in brain tumours has not been extensively studied, but it has a bright future in the coming years. Here, we present a literature review on the role of machine learning in liquid biopsy of brain tumours.

DOI: <https://doi.org/10.47391/JPMA.24-46>

Introduction

Many brain tumours carry significant morbidity for sampling, with some even inaccessible for biopsy. Also, monitoring disease progression and differentiating pseudo-progression requires repeated tissue sampling, which is not feasible.¹ The recent 2021 WHO CNS classification, classifies brain tumours based on molecular characteristics rather than histology. Various markers released by tumour tissue in body fluids can be sampled. Liquid biopsy has remarkable benefits, offering a minimally invasive, safer, faster, repeatable and cheaper way to diagnose and monitor malignant diseases. Samples from CSF, blood (plasma, serum), and urine were used for a liquid biopsy. The tumour-derived material used in the analysis includes circulating tumour cells (CTCs), tumour-educated platelets (TEPs), cell-free DNA (cfDNA), circulating tumour DNA (ctDNA), miRNA, extracellular vesicles and particles (EVPs), and serum-derived small extracellular vesicles (EVs). One of the problems with liquid biopsy is the large amount of data and its complex nature, which can be solved with machine learning (ML). ML algorithms can identify and analyze large complex data sets and build simple models.

Section of Neurosurgery, Department of Surgery, Aga Khan University, Karachi, Pakistan.

Correspondence: Muhammad Shahzad Shamim.

e-mail: shahzad.shamim@aku.edu

ORCID ID: 0000-0001-8305-8854

Various machine learning algorithms used are random forest (RM), support vector machine (SVM), particle swarm optimization (PSO), principal component analysis (PCA), adaptive boosting (AdaBoost), partial least squares-discriminant analysis (PLS-DA), neural networks, decision trees (DT), gradient boosted decision trees, k-nearest neighbor (k-NN), and linear regression (LR).

Literature Search

Morokoff et al., profiled pre- and post-operative serum miRNA in 91 glioma patients and 17 healthy controls using a machine-learning random forest analysis.² They identified a highly accurate 9-gene miRNA signature that could discriminate between glioma and healthy controls with 99.8% accuracy. Two miRNAs, miR-223 and miR-320e, best demonstrated dynamic changes that linked closely with tumour volume in low-grade glioma (LGG) and Glioblastoma, respectively. Most important, miRNA levels did not increase in two cases of pseudo-progression, implying that miRNA is reflective of the volume of glioma cells in the brain and is not influenced by other changes such as inflammation, necrosis or oedema on MRI that are associated with pseudo-progression.

Mouliere et al., recruited 35 glioma patients (30 high grade gliomas (HGG), five LGG), 26 healthy volunteers and 27 patients with other pathologies of the central nervous system and studied cfDNA in CSF, urine and plasma using two sequencing-based approaches, patient-specific hybrid-capture panels and shallow whole genome sequencing (sWGS).³ They found that cfDNA in urine of glioma patients was significantly more fragmented compared to urine from patients with non-malignant brain disorders and healthy individuals. Machine learning models integrating fragment length could differentiate urine samples from glioma patients suggesting possibilities for truly non-invasive tumour diagnosis.

Theakstone et al., studied the role of LB comparing the volume of brain tumour on MRI.⁴ They recruited 90 patients with HGG or LGG with 87 controls. Tumour volumes were calculated from magnetic resonance imaging (MRI) and patients were divided into two groups based on tumour post-contrast enhancement or not. Using attenuated total reflection (ATR)-Fourier transform infrared (FTIR) spectroscopy on serum coupled with supervised learning methods and machine learning algorithms, 90 tumour

patients were correctly classified as either glioma or non-glioma with tumour volumes as small as 0.2 cm³ proving that this technique has great promise for early diagnosis.

Tsvetkov et al. found that the denaturation profile in blood of 84 glioma patients and 63 healthy controls was different with 92% accuracy with help of ML algorithms.⁵ They proposed development of a LB diagnosis program with 2 stages. First artificial intelligence (AI) stage where denaturation profiles of more and more diagnosed glioma patient create an atlas and second blood test stage where denaturation state of untested individual is compared by AI and tells about the disease state of individual.

Bukva et al., studied molecular content of small extracellular vesicles (sEVs) by Raman spectroscopy.⁶ They studied 138 serum samples from four patient groups (glioblastoma, non-small-cell lung cancer brain metastasis, meningioma and lumbar disc herniation as control). They performed the Principal Component Analysis–Support Vector Machine (PCA–SVM) algorithm on the Raman spectra for pairwise classifications and found excellent classification results showing the promising future in Raman spectroscopic analysis for diagnosing and classification of brain tumours.

Sol et al., studied the role of tumour educated platelets (TEP) in diagnosis and monitoring of glioblastoma.⁷ They collected and isolated platelet pellets from whole blood by differential centrifugation from 89 patients with primary glioblastoma on the day of tumour resection (From 52 of these 89 patients, follow-up blood samples were collected during postoperative chemo-and radiotherapy treatment), 126 patients with brain metastases, 86 patients with multiple sclerosis and 353 asymptomatic healthy controls. They found detection accuracy of glioblastoma of 95%. The digital SWARM algorithm demonstrated that the TEP tumour scores of glioblastoma patients could be used to distinguish false positive progression from true progression.

Mikolajewicz et al., performed proteomic profiling of 73 CSF samples obtained from patients with normal pressure hydrocephalus (n=20), glioblastoma (n=22), brain metastasis (n=17), or primary central nervous system lymphomas (n=14) and found Proteomic signatures using machine learning classifiers and survival analysis.⁸ They found 755 unique proteins using 30 µL CSF volumes. Novel biomarkers were identified, including GAP43, TFF3 and CACNA2D2. They concluded that Reliable classification of intra-axial malignancies using low CSF volumes is feasible, allowing for longitudinal tumour surveillance.

Other than glioma, role of ML and LB has also been explored in pituitary tumours. Herrgott et al., studied the

Table: Summary of papers on Machine Learning application to liquid biopsy of brain tumours

S.No.	Study	Year	Medium	Tumour derived material	Sample size	Brain tumour	ML algorithm	Result	Accuracy
1.	Morokof et al. ²	2020	Serum	miRNA	108	Gliomas	Random Forest	9-gene miRNA signature distinguished between glioma and healthy controls.	99.8%
2.	Mouliere et al. ³	2021	CSF, Plasma, Urine	ctDNA	88	High grade and low grade gliomas	Linear Regression, random forest, support vector machine	ctDNA in urine of glioma patients was significantly more fragmented compared to urine from patients with non-malignant brain disorders and healthy individuals	92%
3.	Theakstone et al. ⁴	2021	Serum	ATR-FTIR	177	High grade and low grade glioma	Random forest Support vector machine	Patients with tumour volumes as small as 0.2 cm ³ were correctly identified	88%
4.	Tsvetkov et al. ⁵	2021	Plasma	Denaturation profiles	147	22 oligodendroglioma, 25 astrocytoma, 37 glioblastoma	Random forest Support vector machine, Linear regression	Denaturation profiles could be automatically distinguished	92%
5.	Bukva et al. ⁶	2021	Serum	Small extracellular vesicles (Raman Spectroscopy)	138	Glioblastoma, brain metastases, meningiomas, lumbar disc herniations	Principal component analysis–support vector machine, Free Viz	Groups were distinguishable	82.9–92.5%
6.	Sol et al. ⁷	2022	Blood	Tumour educated platelet RNA	805	Glioblastoma, brain metastases, multiple sclerosis	Particle-swarm intelligence enhanced support vector machine algorithm	Could distinguish false positive progression from true progression	85%
7.	Mikolajewicz et al. ⁸	2022	CSF	Proteomics	73	Glioblastoma, brain metastases, primary CNS lymphoma	Linear regression machine learning classifier framework,	Proteomic-based LR classifiers identified malignancy	94%
8.	Herrgott et al. ⁹	2022	Serum, plasma	cfDNA	100	PIN1ET, non-PIN1ET, other pituitary diseases	Random forest	Could distinguish PIN1ET from non-PIN1ET conditions, including sellar tumours and non-neoplastic pituitary diseases.	93%

methylation-based LB profiling in 59 serum and 41 plasma LB specimens from patients with pituitary neuro-endocrine tumours (PitNETs) and other CNS diseases (sellar tumours and other pituitary non-neoplastic diseases, lower-grade gliomas, and skull-base meningiomas) or non-tumour conditions, grouped as non-PitNET.⁹ They found that methylome landscape of PitNETs was distinct as compared to non-PitNETs with above 93% accuracy. Thus proving the role of LB to diagnose and potentially impact the prognostication and management of patients with PitNETs.

Conclusion

ML has role in LB for early detection of brain tumours, rapid prediction of effectiveness of various therapies and correct differentiation of progression versus pseudo-progression. Currently there is paucity of researches but role of ML in LB of brain tumours excites many researchers and may have future in brain tumour management.

References

1. Menna G, Piaser Guerrato G, Bilgin L, Ceccarelli GM, Olivi A, Della Pepa GM. Is There a Role for Machine Learning in Liquid Biopsy for Brain Tumors? A Systematic Review. *Int J Mol Sci.* 2023;24:9723. doi: 10.3390/ijms24119723. PMID: 37298673; PMCID: PMC10253654.
2. Morokoff A, Jones J, Nguyen H, Ma C, Lasocki A, Gaillard F et al. Serum microRNA is a biomarker for post-operative monitoring in glioma. *J Neurooncol.* 2020;149:391-400.
3. Mouliere F, Smith CG, Heider K, Su J, van der Pol Y, Thompson M et al. Fragmentation patterns and personalized sequencing of cell-free DNA in urine and plasma of glioma patients. *EMBO Mol Med.* 2021;13:e12881.
4. Theakstone AG, Brennan PM, Jenkinson MD, Mills SJ, Syed K, Rinaldi C et al. Rapid spectroscopic liquid biopsy for the universal detection of brain tumours. *Cancers (Basel).* 2021;13:3851. doi: 10.3390/cancers13153851. PMID: 34359751; PMCID: PMC8345395.
5. Tsvetkov PO, Eyraud R, Ayache S, Bougaev AA, Malesinski S, Benazha H et al. An AI-powered blood test to detect cancer using nanoDSF. *Cancers.* 2021;13:1294.
6. Bukva M, Dobra G, Gomez-Perez J, Koos K, Harmati M, Gyukity-Sebestyen E, Biro T, Jenei A, Korondi S, Horvath P, Konya Z, Klekner A, Buzas K. Raman Spectral Signatures of Serum-Derived Extracellular Vesicle-Enriched Isolates May Support the Diagnosis of CNS Tumors. *Cancers (Basel).* 2021;13:1407. doi: 10.3390/cancers13061407. PMID: 33808766; PMCID: PMC8003579.
7. GJG S, Vancura A, Tjerkstra M, Leurs C, Rustenburg F et al. Tumor-educated platelet RNA for the detection and (pseudo) progression monitoring of glioblastoma. *Cell Rep Med.* 2020;1: 100101. doi: 10.1016/j.xcrm.2020.100101. PMID: 33103128; PMCID: PMC7576690.
8. Mikolajewicz N, Khan S, Trifoi M, Skakdoub A, Ignatchenko V, Mansouri S et al. Leveraging the CSF proteome toward minimally-invasive diagnostics surveillance of brain malignancies. *Neurooncol Adv.* 2022;4:vdac161.
9. Herrgott GA, Asmaro KP, Wells M, Sabedot TS, Malta TM, Mosella MS et al. Detection of tumor-specific DNA methylation markers in the blood of patients with pituitary neuroendocrine tumors. *Neurooncol.* 2022;24:1126-39.