RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt2665

Analysis of Gene Co-Expression Networks using Centrality Measures in order to Identify Significant Genes Correlated with Cancer Diagnosis

Seema Sandeep Redekar [a, *], Smita Dange [b], Madhura Vyawahare [c]

[a] Department of Information Technology, K. J. Somaiya School of Engineering (formerly known as K J Somaiya College of Engineering), Somaiya Vidyavihar University, Mumbai, 400077, India

[b] Fr.C.Rodrigues Institute of Technology, Vashi, Navi Mumbai, India

[c] SVKM's NMIMS deemed to be University, Navi Mumbai, India

* Corresponding Author Email: seema.r@somaiya.edu

DOI: https://doi.org/10.54392/irjmt2665

Received: 09-03-2026; Revised: 15-09-2026; Accepted: 21-10-2026; Published: 30-10-2026

Abstract: One of the most aggressive primary brain tumors is glioblastoma multiforme (GBM), which has poor prognosis, and with limited therapeutic options. Identifying reliable molecular markers will be crucial to enhancing the prognostic value for GBM and to elucidating the biological mechanisms of GBM progression. While co-expression network analysis has been used extensively for biomarker discovery, traditional centrality measures in graph theory assign only topological measures to genes, and exclude any biological information based on patient survival. In this study, we suggest a biologically informed node-weighted centrality approach for incorporating survival-related biological information with the conventional network topology to enhance biomarker prioritization. This dataset contains DNA methylation profiles from the The Cancer Genome Atlas (TCGA) GBM cohort which was preprocessed to bring out informative probes followed by univariate Cox proportional hazards analysis to extract the genes associated with survival. To incorporate biological weights that are derived from survival, node-weighted versions of degree, closeness, betweenness and eigenvector centrality were formulated on the biological network. The proposed framework was then used to prioritize genes within the co-expression network that have a biological relevance. A total of 156 genes were identified in the analysis that are associated with survival, and ten candidate prognostic biomarkers (SEMA3B, MARK2, APS, PITPNM2, SFRP1, SLC44A2, PRLH, KRTAP4.2, CTSZ and DIP2C) were selected. These biomarkers were uniformly associated with overall survival and are biologically meaningful molecular signatures with survival data and network topology. This study introduces a biologically informed node-weighted centrality framework which integrates the biologically derived information from the survival analysis into the graph-theoretic network analysis for biomarker prioritization. The proposed methodology is an extension of our previously reported biomarker discovery pipeline, where the node weighting is introduced using a survival-guided approach, while the previously validated biological dataset and the biomarker set remain. This framework offers a computational approach to biomarker prioritization based on networks of biological relevance which can be extended to GBM and other complex diseases.

Keywords: Gene Co-Expression Network, Centrality Measures, Weights, Survival Analysis, Node Weighted Centrality.

1. Introduction

Cancer is often described as a “genetic disease” because it occurs when alterations in the sequence of genes that control cell growth, differentiation, and normal cellular functions occur. There are a group of genes the function of which is critical and when a series of mutations build up over time, this causes the cells to grow out of control and become cancerous [1]. There are also epigenetic changes in gene expression that do not alter the DNA sequence – including DNA methylation, histone modification and chromatin remodeling, which also contribute to the progression of cancer. In this regard, DNA methylation has been identified as one of the most relevant molecular processes linked to tumor initiation, progression, and prognosis in patients, which makes it a promising source for the discovery of biomarkers in precision oncology [2, 3].

Among the numerous types of cancer, glioblastoma multiforme (GBM) is one of the most

aggressive tumors of the central nervous system [4]. Although progress has been achieved in surgery, radiotherapy and chemotherapy, GBM is known for its aggressive growth, diffuse infiltration, great intratumoral heterogeneity and resistance to standard treatment modalities, leading to poor clinical outcomes [5, 6]. Patients with GBM have a median survival of 12–15 months and 5-7 year survival is less than 10%, making the need for more accurate molecular markers to enhance the ability to predict patient survival and guide personalized therapeutics highly urgent [7, 8]. Thus, the efficient computational approaches to understand GBM progression and finding predictive markers for GBM prognosis are still important research tasks.

Multiple genes and regulatory pathways are involved in the progression of GBM, which interact in a complex manner [9, 10]. Therefore, knowledge of these molecular interactions is crucial for unveiling disease mechanisms and finding clinically relevant biomarkers. The application of new high-throughput sequencing technologies and large-scale cancer repositories like The Cancer Genome Atlas (TCGA) has allowed for the study of genome-wide molecular changes including DNA methylation, copy number variation, gene expression and mutation profiles. Large-scale data of these molecular features and computational network analysis has greatly helped in identifying candidate biomarkers related to patient survival and disease progression [11, 12].

Biological processes can be represented in an intuitive way through biological networks, as they represent molecular interactions among genes, proteins and other biological entities [13]. The network can be directed or undirected, depending on the model of the biological relationship. The gene co-expression relationships that are derived from DNA methylation profiles are represented as an undirected graph in this study [14]. Gene co-expression networks (GCNs) link genes that are statistically significantly co-expressed and have proven useful in studying the relationship among genes and for finding modules involved in diseases [15, 16]. Network-based approaches capture coordinated biological behavior and could enable identification of highly connected or influential genes that may play key roles in disease development, unlike analyses of individual genes.

Graph-theoretic measures can be used to quantify the structural importance of a node in a biological network [18] and thus be used to investigate the topology of these networks. Of these measures, degree, closeness, betweenness, and eigenvector centrality are commonly used for finding influential genes in a biological network [19, 20]. Degree centrality, which reflects local connectivity, is the indicator of communication efficiency within the network; betweenness centrality, which reflects genes that link functional modules, is an indicator of communication efficiency among genes; eigenvector centrality quantifies the influence of a gene based on the importance of the neighboring genes [21]. These traditional centrality measures have been used to successfully identify biomarkers in networks, but they are largely based on network topology and tend to not rely on biological data, like patient survival or clinical significance, in their measure of centrality. This means that relatively "modest" topological characteristics are not sufficient to make a biologically important gene stand out in network analysis [22, 23].

A key statistical tool for assessing the association between molecular biomarkers and patient outcomes is survival analysis. One of the most popular approaches to determining genes associated with survival is the Cox proportional hazards model that can estimate hazard ratios and significance values for those genes while accommodating censored observations [24]. Genes that have important implications for overall survival are logical candidates for the development of prognostic biomarkers. Most of the time, however, survival information is only applied as a filtering process in the construction of the network, but it is not applied directly to network topologies analysis [25]. This constraint drives the design of biologically-inspired approaches to network analysis which incorporate biological information from the longevity data onto the graph structure. Adding the statistical evidence derived from survival analysis into the centrality calculation can better reflect both the structural importance of candidate biomarkers in the biological network and the clinical significance. This can lead to better biological interpretability of network-based biomarker prioritization without changing the structure of the network [26].

This current study extends our previous study [1] to explore the gene co-expression network analysis of glioblastoma multiforme (GBM) based on TCGA DNA methylation data. The previous studies were limited to the identification of the prognostic biomarkers by applying traditional graph theoretical centrality. For this reason, the cohort used in the current study is the same as that used in our previous study, the processing pipeline, the probe filtering, the gene mapping, and the survival-associated gene selection are the same as used in our previous study, and we have validated all these pipelines. The aim of this reuse was to enable a consistent methodological approach and the assessment of the proposed computational framework on the same biological basis. The main feature of this work is the development of a biologically informed node-weighted centrality framework that integrates survival-derived biological information with traditional centrality computation to prioritize biomarkers based on network topology and biological significance. Therefore, the biological data set and the proposed biomarker set are the same, and the novelty is the proposed weighted network analysis framework. This study is not a standalone biomarker discovery study, but rather a clear

methodological extension of our previous reported biomarker discovery pipeline.

This study's findings can be summarized as follows:

- Development of a node-weighted centrality framework based on biological information obtained through survival analysis and graph-theoretic centrality measures.
- Development of node-weighted degree, closeness, betweenness, and eigenvector centrality in biological networks.
- Analysis of proposed framework for prioritisation of survival-associated biomarkers based on TCGA GBM DNA methylation data.

A clear distinction between the re-used biological analyses and novel contribution of the computational analysis, and a transparent methodological comparison with our previous study is provided in this study.

All newly generated figures and newly visualized/reproduced/adapted figures from our previous publication (with proper citations) have been clearly marked in the revised manuscript to ensure complete transparency.

2. Literature Review

The discovery of prognostic biomarkers for glioblastoma multiforme (GBM) has greatly benefited from recent developments in high throughput sequencing technologies, large scale cancer genomic repositories, and computational network analysis. Many studies have examined DNA methylation profiling, gene co-expression networks, survival analysis and graph-theoretic approaches, among others, for biomarker discovery. Current approaches, however, vary widely in data sources, analytical approaches, and priorities for biomarkers, which can vary in their biological interpretability and clinical applicability. Accordingly, this section critically examines recent literature to pinpoint the limitations and the research gap that the proposed framework of biologically informed node-weighted centrality fills.

2.1 Molecular Biomarkers and DNA Methylation in Glioblastoma

Glioblastoma multiforme (GBM) is the most malignant primary brain tumor, with a rapid growth rate, high molecular heterogeneity and unfavorable clinical outcome even with the latest advances in surgical, radiation and chemotherapy treatments. Molecular features have been incorporated into the WHO 2021 Classification of Central Nervous System Tumors, and have enhanced the diagnosis and classification of GBM by combining biomarkers with conventional histopathology [17, 18]. However, the complex biology of GBM requires the discovery of other biomarkers that will help to better stratify patients and guide personalized treatment [19].

DNA methylation is one of the epigenetic mechanisms that have been found to be a good biomarker for cancer diagnosis and prognosis. Tumors develop through aberrant methylation which can silence tumor suppressor genes or activate oncogenic pathways, which can lead to the initiation and progression of tumor development [20]. With the help of resources like The Cancer Genome Atlas (TCGA), genome-wide methylation profiling has helped identify genes and molecular signatures associated with survival in GBM, as well as those involved in GBM development [21].

Recently, other biomarkers have been associated with tumor proliferation, invasion, angiogenesis and therapeutic resistance, such as IDH1, EGFR, TP53, ATRX and PDGFRA [22]. Such investigations are based on testing individual genes or isolated molecular changes, however, the progression of GBMs involves complex interactions between many genes and regulatory pathways. Thus, genetic studies of a single gene may not reveal the coordinated biological processes responsible for tumor progression. Owing to this drawback, network-based methods have gained more and more interest in understanding the molecular interactions at the genome-wide level.

Biological networks can be used to find genes which are functionally important by connecting them together and analyzing them with computation, not just looking at them individually. Survival analysis using Cox proportional hazards model is a common statistical method for discovery of prognosis genes associated with patient survival [23]. However, most studies so far have used survival data as a pre-processing step before building the network and seldom directly in the network topology when prioritizing biomarkers. This constraint is a significant research gap. Current graph-theoretic approaches focus on the genes mostly based on topological features rather than biological significance derived from survival.

Thus, computational approaches that combine network topology and biological information are needed to undertake more biologically relevant biomarker prioritization. The present study aims to fill this gap by developing a biologically-informed node-weighted centrality framework which integrates survival-related biological information into the traditional graph-theoretic centrality measures. This work is made as part of the corrections suggested by the reviewers and is presented as a methodological extension of the previously published biomarker discovery pipeline, instead of a new study of biomarker discovery.

2.2 Gene Co-expression Networks for Biomarker Discovery

Gene co-expression networks (GCNs) have emerged as one of the major systems biology tools to explore gene functional relationships and to discover disease-related biomarkers. In a GCN, genes showing similar expression profiles in various biological samples are linked and the genes and gene modules that co-regulate can be identified, and their role in the disease process can be understood [24]. Network analysis methods over the conventional differential gene expression analysis are able to capture coordinated biological interactions and give a comprehensive understanding of complex diseases like GBM. One of the most popular methods to identify gene modules related to clinical traits is weighted Gene Co-expression Network Analysis (WGCNA), which has been used for tumor progression and patient survival [25, 26]. Various studies have demonstrated that the hub genes of co-expression modules are highly linked to GBM progression and immune regulation, angiogenesis and therapeutic resistance [27]. In addition, incorporating transcriptomic, epigenetic and clinical data has increased the biological significance of the network-based biomarkers and increased the classification of diseases [28].

However, the current GCN-based research focuses mainly on the identification of hub genes based on traditional centrality measures, such as degree, betweenness, closeness, and eigenvector centrality. The approaches focus on the topology of a network and assign a rank to genes based solely on their position in the network, with the hope that highly connected genes are biologically significant [29]. But the clinical significance doesn't necessarily reflect network topology. Genes showing high association with patient survival may be given lower ranks due to having relatively low connectivity in their networks, whereas highly connected genes without prognostic value may be given higher ranks.

A limitation of this approach is that the biological network is not weighted, so recent studies have investigated nodes or edges' attributes based on gene expression, protein interaction or functional annotations [30]. However, only a few studies directly combine survival derived biological information with graph-theoretic centrality computation for prioritizing biomarkers. As a result, computational models that take both of these factors into account continue to be needed.

To address this research gap, the present study introduces a biologically-informed node-weighted centrality framework in which the biologically-informed information from the survival is added to the conventional centrality measures. The proposed framework is biologically informed for biomarker prioritization compared to existing approaches based on topology, both in terms of its importance for the structure of the co-expression network as well as its prognostic relevance.

2.3 Survival-Based Biomarker Identification

Survival analysis is an essential tool in cancer biomarker discovery, which examine the relationship between molecular changes and survival. The Cox proportional hazards (Cox-PH) model is the most commonly used approach to identification of prognostic genes as it can quantify the association between the expression of the gene and overall survival, even when the data is censored. Kaplan–Meier (KM) survival analysis is a frequently used method to plot survival differences between patient subgroups and to test the prognostic value of candidate biomarkers [19].

In recent years, some researchers have used the survival analysis method to discover prognosis markers from a large number of samples such as The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) in glioblastoma.Recently, a number of studies have applied survival analysis to large-scale datasets, including The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), to identify prognosis markers in glioblastoma. These studies combine survival information with transcriptomic and epigenetic analysis to identify genes linked to tumor progression, therapeutic responses, and prognosis of patients [22].

Even with these developments, however, survival analysis is typically considered as a stand-alone statistical filtering process before downstream analyses, such as the construction of networks and identification of hub genes. As a result, the information contained in the Cox regression model is seldom directly used in graph-theoretic network analysis to enhance the biological relevance of the topologies used for biomarker prioritization [31]. Current network studies mostly focus only on structural connectivity and the survival-based biological significance of genes is ignored.

This restriction encourages the creation of biologically inspired computational models that incorporate survival data in network topology. To fill this gap, the proposed node-weighted centrality framework integrates the biological information of survival into the traditional centrality measures to rank biomarkers based on both topologic significance and their prognosis importance in the gene co-expression network.

2.4 Biological networks and graph theoretic/weighted centrality methods

The graph-theoretic methods have been gaining relevance in the context of biological network analysis in order to identify influential genes linked to disease development. Gene co-expression networks are constructed, where the genes are nodes and the

interactions between them are edges, and were used to identify biologically important genes through the use of graph-theoretic centrality measures [32]. A variety of centrality measures have been developed and used, including degree centrality, which identifies genes with high connectivity; closeness centrality, which assesses information flow efficiency; betweenness centrality, which identifies genes as communication hubs; and eigenvector centrality, which evaluates the influence a gene has on the importance of its neighbors.

These centrality measures have been used extensively to pinpoint hub genes in several types of cancer such as GBM, assuming that nodes with high centrality have important biological functions [33]. Conventional centrality measures, however, are based purely on network topology and typically are not influenced by biological evidence, like survival significance, gene function or clinical relevance, in the process of ranking genes. Several researchers have suggested weighted biological networks by assigning weights to the nodes or edges, which are based on the gene expression level, protein interaction, functional annotation, or pathway information [34]. While these methods enhance the representation of networks, few studies have been conducted to incorporate the information in a biological system in the form of survival into the computation of centrality in a graph for the purpose of prioritising biomarkers. Therefore, if genes have weak positions with regards to the network positions, they may not get higher ranks if they have strong prognostic significance.

In the current study, this limitation is overcome by proposing a biologically informed node-weighted centrality framework which embeds the survival-derived biological information into the traditional centrality measures. The proposed framework is different from the topology-based methods, as it accounts also for the biological significance of the network while prioritizing the biomarkers. To facilitate the rest of this manuscript, the modifications applied to this manuscript are applied to this work as well, with the novelty coming in the proposed weighted centrality formulation and not from the biological dataset or biomarker set used.

2.5 Critical Analysis and Research Gap

Despite the success of the molecular characteristics of glioblastoma identified through molecular profiling, gene co-expression analysis, survival modelling and graph-theoretic approaches, there are methodological drawbacks. Although recent studies have shown that DNA methylation profiling and multi-omics integration can be used to identify prognosis related biomarkers, their studies are mainly based on statistical significance or differential expression analysis without considering the structural importance of genes in biological interaction networks [35, 36]. In contrast, the gene co-expression network studies rely on centrality of the genes, which are traditionally computed using graph-theoretic algorithms, and generally focus only on the topology of the network and not on clinically relevant biological signals such as the survival of patients [37].

Likewise, to better represent a biological network, weighted biological network approaches have incorporated node or edge weights based on either gene expression or protein interactions or functional annotations [3]. However, few research efforts have incorporated directly into the centrality computation for biomarker prioritisation information derived from the survival outcome of clinical trials. This means that genes that have high topological connectivity but low prognostic relevance can be ranked lower than genes that may have low topological connectivity but high clinical relevance.

A major drawback of recent studies is that the focus has been on discovery of new biomarker sets with only marginal attention on the improvement of the computational methods applied to biomarker prioritisation. Thus there is a lack of development of methodological innovations that make the biological interpretation of network analysis more feasible.

To fill this gap, the present study introduces a biologically informed node-weighted centrality framework that combines with graph-theoretic centrality measures the "survival" information derived from biological data. The proposed approach takes network topology and prognostic relevance into account as a single entity, unlike existing approaches. In keeping with the changes proposed by the reviewers, the current article is actually stated as a methodological extension of our previous article that reported discovery of a new biomarker set, but rather the discovery of a new biomarker discovery pipeline. The novelty is therefore the proposed node-weighted centrality formulation that not only improves the biological interpretation of the network-based biomarker prioritization methodology but also keep it fully transparent. The comparative analysis of the papers studied to understand biomarkers discovery linked with various cancer is presented in Table 1.

It is clear from the above survey that current studies either use survival analysis as a statistical filter, or use the graph-theoretic centrality measure based on only the topological structure of the network. Only a handful of methods incorporate directly survival information into centrality measures to prioritize biomarkers. The methodological void inspires the present study to propose a biologically-informed node-weighted centrality framework that integrates the prognostic significance with network topology. Most importantly this work proposes an extension of our previously published biomarker discovery pipeline in a transparent and reproducible way, instead of proposing a new biomarker set.

Table 1. Comparative Analysis of Recent GBM Biomarker Discovery Studies

Ref.	Study	Dataset	Methodology	Survival Analysis	Network Analysis	Limitation
Louis *et al.* (2021) [17]	WHO CNS Classification	Multiple cohorts	Molecular classification	No	No	Diagnostic focus only
Lan *et al.* (2024) [36]	GBM biomarkers review	Literature	Molecular biomarker review	Yes	No	No network prioritization
Shah *et al.* (2024) [19]	Prognostic biomarkers	TCGA	Multi-biomarker analysis	Yes	No	Individual biomarker analysis
Jacome *et al.* (2024) [22]	Precision medicine	TCGA + Literature	Multi-omics	Yes	Limited	No topology integration
Zhang *et al.* (2023) [25]	WGCNA	TCGA	Gene co-expression	Yes	Yes	Conventional centrality
Li *et al.* (2022) [35]	Integrated transcriptomics	TCGA	Co-expression modules	Yes	Yes	No biological node weighting
Wang *et al.* (2023) [27]	Hub gene analysis	TCGA	Degree centrality	Yes	Yes	Topology only
Chen *et al.* (2024) [33]	Multi-omics network	TCGA	Network integration	Yes	Yes	Survival not integrated into centrality
Ma *et al.* (2023) [38]	Weighted biological n/w	Cancer datasets	Weighted network	Limited	Yes	Weight not derived from survival

3. Materials and Methods

This study is based on the overall flow of the work as shown in Figure 2. The proposed approach consists of four steps: (i) DNA methylation dataset pre-processing and acquisition, (ii) identification of survival associated genes from previously validated results of survival analysis, (iii) construction of the network of the co-expressed genes and (iv) identification of key biomarkers using proposed biologically informed node-weighted centrality framework. The dataset preparation and the selection of survival-associated genes were intentionally reused from our previous study [1] as the aim of the work was not to develop a new preprocessing pipeline but rather adopt the proposed weighting strategy. The node-weighted centrality formulation and the network-based biomarker prioritization are the methodological contributions of the present work.

3.1 Dataset Description and Study Design

The current study took advantage of the Glioblastoma Multiforme (GBM) DNA methylation dataset from The Cancer Genome Atlas (TCGA), a publicly accessible cancer genomics data repository that provides extensive genome-wide molecular and clinical information for various cancer types [39]. The TCGA GBM DNA methylation dataset was chosen as it contains genome-wide methylation data with associated clinical and survival data, allowing for the investigation of survival associated biomarkers using network analysis. This work is an extension of the study, which we published earlier [1], in the methodological part. Thus, the same TCGA GBM DNA methylation dataset, preprocessing and probe filtering, gene annotation and identification of survival-associated genes were intentionally reused to offer a validated biological basis for assessing the proposed node-weighted centrality framework. The aim of the present study is not to develop a new set of biomarkers, but rather to explore if using survival derived biological information in graph-theoretic centrality measures offers a biologically informed framework for prioritizing biomarkers. Therefore, the preprocessing procedure and the set of reported genes associated with survival described in our previous publication was not changed.

3.2 Data Preprocessing

This is the original TCGA GBM DNA methylation dataset consisting of 288 patient samples. This is the original TCGA GBM DNA methylation dataset, with 288 patient samples. In the preprocessing phase, patients

with incomplete methylation data for the probes selected were excluded. Samples with incomplete survival data (where overall survival time or survival status was missing) were subsequently excluded as these data are needed for Cox proportional hazards regression and Kaplan–Meier survival analysis. After this pre-processing, 76 patients with complete DNA methylation probes 24,925 and survival data were used for survival analysis as presented in Table 2. The step followed to obtain 76 patients from 288 patient's sample is presented below in Figure 1. The whole preprocessing process and patient selection protocol are shown in Figure 2. After a series of preprocessing, 156 genes associated with survival were identified and these genes were employed in the construction of gene co-expression network and node-weighted centrality analysis.

3.3 Survival Analysis and Gene Selection

The DNA methylation data for gene expression and clinical survival information were acquired from TCGA GBM dataset. The genes that were considered significantly associated with overall survival were identified using survival analysis after preprocessing. Molecular features and patient survival outcomes can be estimated through statistical modelling, known as 'survival analysis' or 'time-to-event analysis' [40]. In the previous study [1] we used univariate Cox proportional hazards regression to determine which genes in the filtered methylation data set were associated with survival. In our previous publication [1] a univariate Cox proportional hazards regression analysis was performed on 11,466 probes that were filtered to yield 156 survival-associated genes [41, 42].

Table 2. The number of genes connected to survival outcomes found by survival analysis to build and analyze gene co-expression communities

Type of Cancer	TCGA Datasets	No of Patients	Probes/ Genes	Significant genes identified using survival analysis
Glioblastoma Multiforme (GBM)	DNA Methylation	288	24,925	156

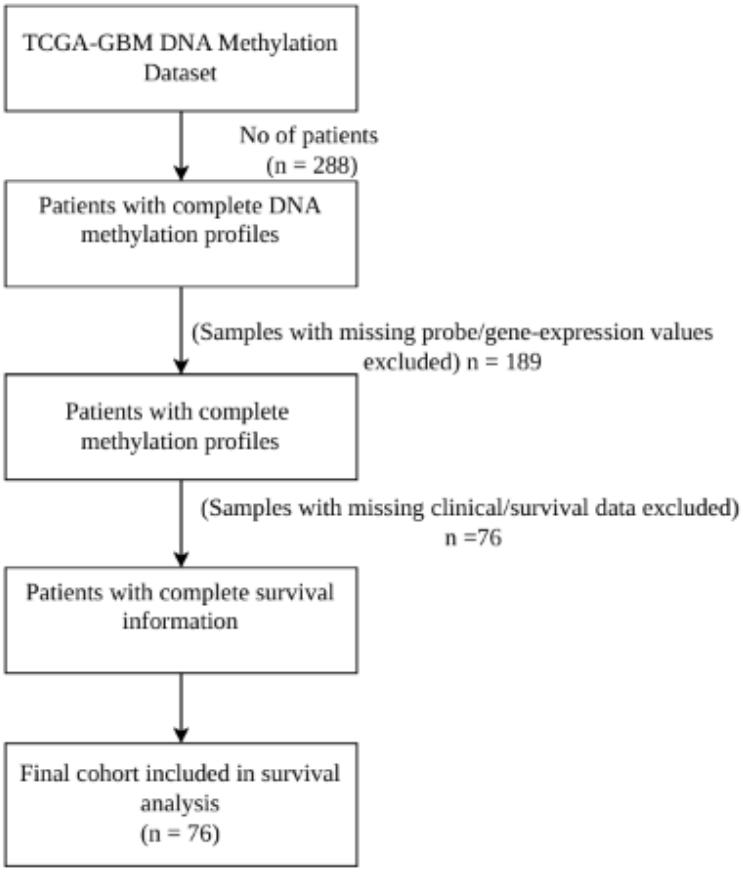


Figure 1. Steps to followed to obtain 76 patients from 288 patients DNA Methylation dataset.

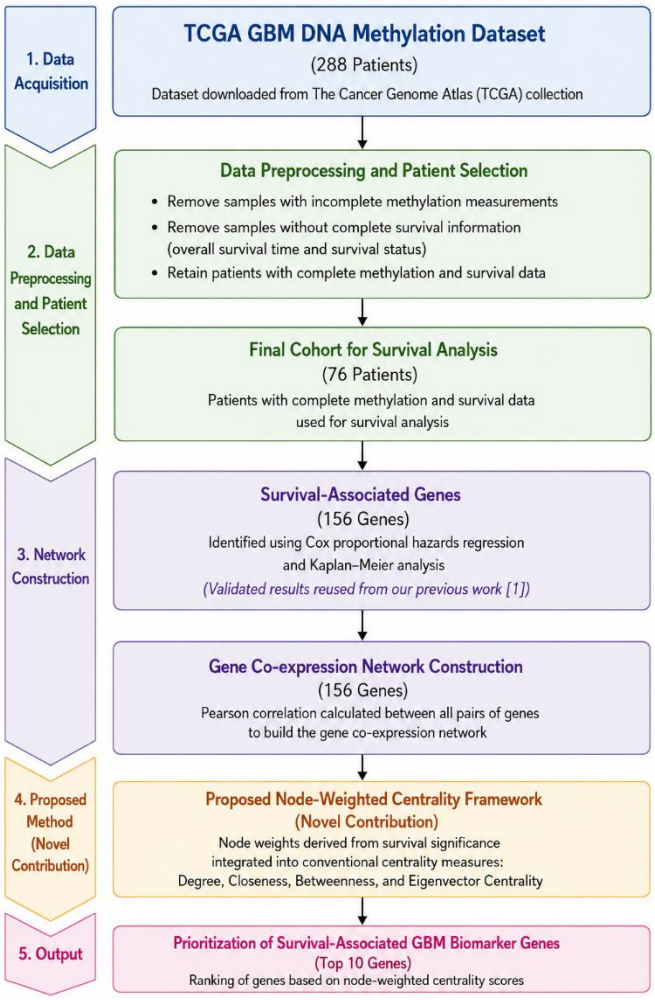


Figure 2. Overall workflow of the proposed node-weighted centrality framework and patient selection process.

For this study, it is important to note that these hazard ratios and associated p-values were intentionally used again as biologically meaningful inputs into the proposed node-weighted centrality framework. So, no new survival analysis was carried out, and the previous survival statistics were used in the node-weight computation.

3.4 Gene Co-expression Network Construction

To build the gene co-expression network, 156 genes associated with survival, obtained by the previously-validated univariate Cox proportional hazards regression analysis [43, 45], were used. After pre-processing, 76 patients with full DNA methylation and survival data were analyzed using the network.

The gene co-expression network is shown as an undirected weighted graph G= (V, E, δ), where V are the genes (nodes), E are the co-expression relationships (edges), and δ are the Pearson correlation coefficient between gene pairs [45, 46]. Pearson correlation coefficients were computed between pairs of genes and an adjacency graph was created by connecting pairs of genes whose absolute correlation coefficient is above a predefined value [47, 48].

Correlation measures of Pearson, Spearman and Kendall were first examined. Since DNA methylation values are continuous variables, Pearson correlation was chosen as the correlation method, as it can effectively capture the linear relationships and offer biologically interpretable gene co-expression networks. A threshold of 0.5 for absolute Pearson correlation was used to create the adjacency matrix. This threshold is selected to meet the consideration of network sparsity and biological connectivity, while retaining moderate-to-strong gene associations while suppressing weak correlation. While the sensitivity analysis with multiple thresholds is not presented here, this is mentioned in the Discussion of the present study. Overall construction process of the network is depicted in Figure 3; Figure 3(a) depicts the gene methylation matrix, Figure 3(b) the Pearson correlation matrix, Figure 3(c) the corresponding adjacency matrix, and Figure 3(d) the gene co-expression network which was employed for the node-weighted centrality analysis.

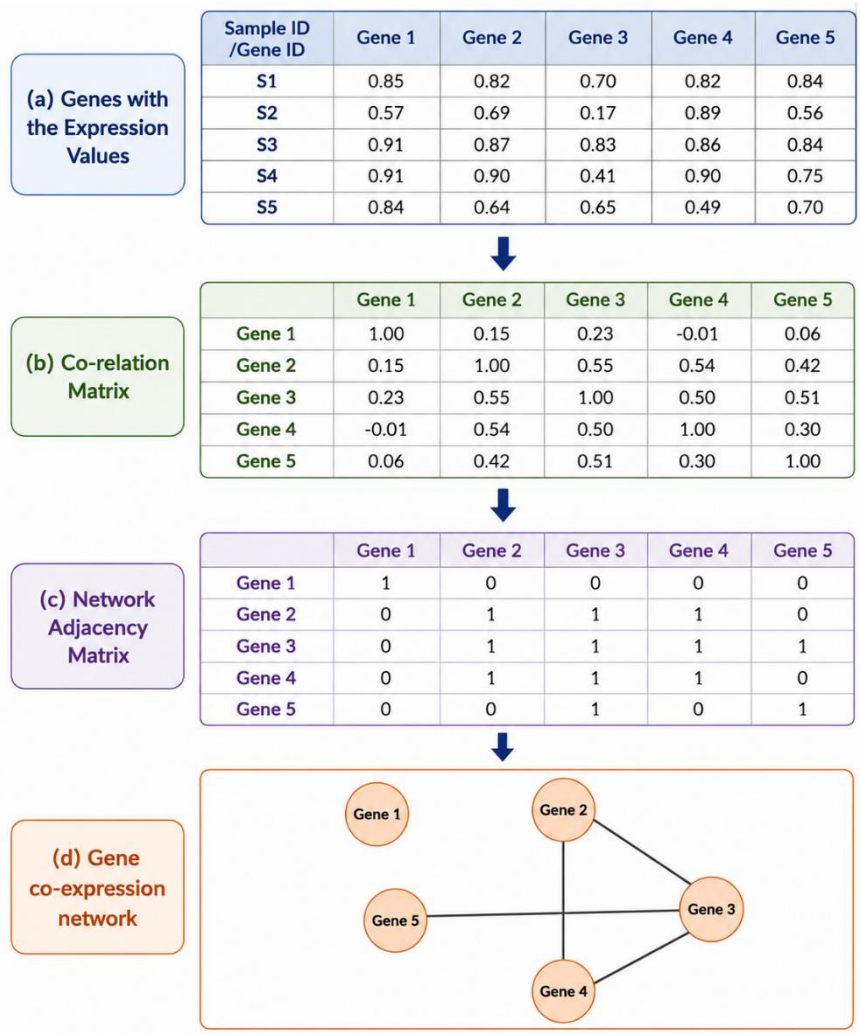


Figure 3. Steps for building a gene co-expression network: **(a)** Genes with expression values **(b)** Co-relation Matrix **(c)** Network adjacency Matrix **(d)** Constructed network (Redrawn and updated from the workflow presented in our previous publication [1] for improved clarity).

3.5 Proposed Node-Weighted Centrality Framework

The objective of the proposed node-weighted framework was not to identify a different biomarker set but to enhance conventional network analysis by incorporating survival-derived biological information into the computation of node importance.

The idea of proposed framework is to improve the traditional graph-theoretic centrality measures by adding the biological information provided by the survival data to the calculation of the node's importance. The aim is not to find a different marker set, but to prioritize the biomarkers by combining the markers' importance in the network with their prognostic value. This biologically informed weighting strategy helps achieve a more clinically relevant weighting of the node depicting genes in the network.

The current study is a methodological continuation of the previous one published [1]. The TCGA GBM DNA methylation dataset was intentionally reused to assure consistency of method used, as were the same preprocessing pipeline and validated survival statistics, along with the same survival-associated gene selection. A novelty of the present study comes after the construction of the gene co-expression network, upon using node-specific weighting of conventional centrality measures to include biological information derived from survival.

In the proposed framework, each node is assigned a biological weight derived from the statistical significance obtained through Cox proportional hazards regression. These node weights are incorporated into Degree, Closeness, Betweenness, and Eigenvector centrality computations, resulting in biologically informed centrality measures for biomarker prioritization. The mathematical formulation of the proposed node-weighted centrality measures is presented below [49].

The weight of each node here is the raw p-value from the univariate Cox proportional hazards regression results reported previously. These p-values were

included in the calculation of the node-weighted centrality measures to incorporate biological information from survival analysis into network topology. The weighting algorithm adds relative survival information to the statistically significant genes, as only genes with p value at or below the predefined significance threshold ($p \leq 0.01$) were included. These weights are applied on the nodes to compute various centrality measures. Degree, Closeness, Betweenness, and Eigenvector Centrality are different centrality measurements [50]. The value for weight is computed in the second step of methodology shown in Figure 1. In the survival analysis phase we have extricated survival associated genes (p-value<=0.01) by applying Cox regression model. This significance value is used as weight while computing node weight in the centrality computation.

$$w_i = p_i \quad (1)$$

Where,

w_i = node weight,

p_i = p-value from the univariate Cox regression.

Newly formulated centrality measures are given below:

3.5.1 Node-Weighted Degree Centrality

Degree centrality is redefined using node weights and with hazard function of Cox model to accommodate the node weights which is a significance value obtained in the summary of Cox model. Following that, node-weighted centrality of u is calculated:

$$DC(u) = \left(\sum_{v \in V \setminus \{u\}} a_{uv}\right) * w_i \quad (2)$$

Where,

w_i : node weight (raw p-value)

a_{uv}: adjacency matrix element

3.5.2 Node-Weighted Closeness Centrality

Closeness centrality is also reformulated with node weight value obtained through hazard function and computed for a node u in a network is as:

$$\mathrm{CC(u)} = \left(\frac{1}{\sum_{v \in V \setminus \{u\}} d_{uv}}\right) * w_i \quad (3)$$

where,

w_i : node weight (raw p-value)

d_{uv}: Shortest-path distance

3.5.3 Node-Weighted Betweenness Centrality

When computing the betweenness centrality using the core formula, it indicates that the value of communication between two pairs is multiplied. The node-weighted betweenness centrality of a node u is obtained through:

$$BC(u) = \left(\sum_{s,t,u \in V:|\{s,t,v\}|=3} \frac{\sigma_{st}(u)}{\sigma_{st}}\right) * w_i \quad (4)$$

Where,

w_i : Node weight (raw p-value)

d_{uv}: Shortest-path distance

σ_{st}: Number of shortest paths

$\sigma_{st}(u)$): number of shortest paths passing through node u

3.5.4 Node-Weighted Eigenvector Centrality

The node-weighted eigenvector centrality of a node u was calculated as:

$$EC(u) = \left(\sum_{v \in V \setminus \{u\}} \left(a_{uv} \ . EC(v)\right)\right) * w_i) \quad (5)$$

Where,

w_i : Node weight (raw p-value)

a_{uv}: Adjacency matrix element

Centrality measures considering weights on node are computed as per the Equation (II)

For node weighted degree centrality, Equation (III) for node weighted closeness centrality,

Equation (IV) for node weighted betweenness centrality and Equation (V) for node weighted eigenvector centrality.

Raw p-values were directly used as the node weights, so genes with smaller raw p-values are scaled proportionally by less. This is a reflection of the weighting used in the present study. Other transformations can be used as long as they are equivalent to using a log transformation, for example, −log10 (p). The network constructed using Pearson correlation measures as given in methodology for 25 nodes (genes with gene id's) is shown Figure 4.

The network shown in Figure 4 is analyzed using node weighted centrality measures [51, 52]. The resultant graph after applying node weighted degree centrality is shown in Figure 5. Figure 5(a) shows nodes with greater node weighted degree centrality values in red and varying sizes. The graph shown in Figure 5(b) is the result of applying node weighted betweenness centrality given in Equation (III). The nodes shown in pink color are the nodes having high node weighted closeness centrality values as per Figure 5(c). The result of application of node weighted eigenvector centrality values are shown with blue colour nodes in Figure 5(d). These centrality measurements are then performed separately to each gene co-expression network. The intersection of all these measures is computed find most important node in the biological network which inferred as most important survival associated genes.

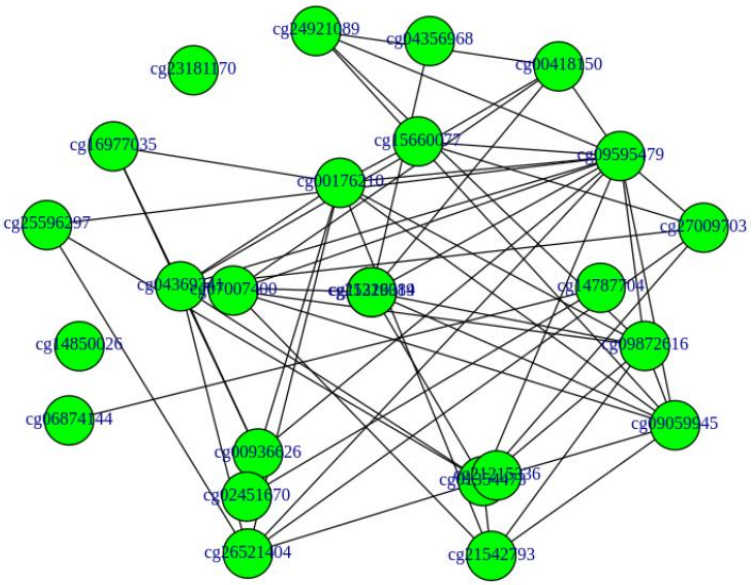
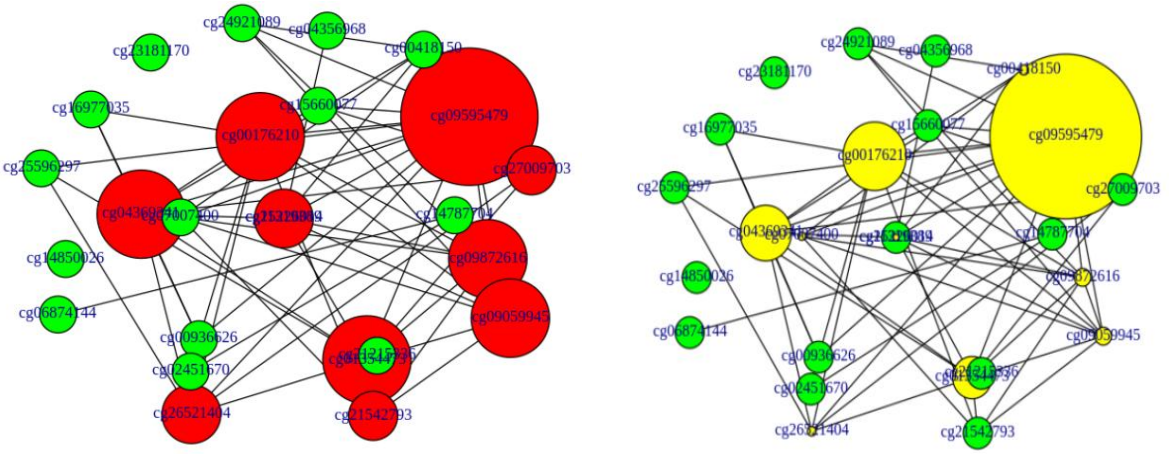
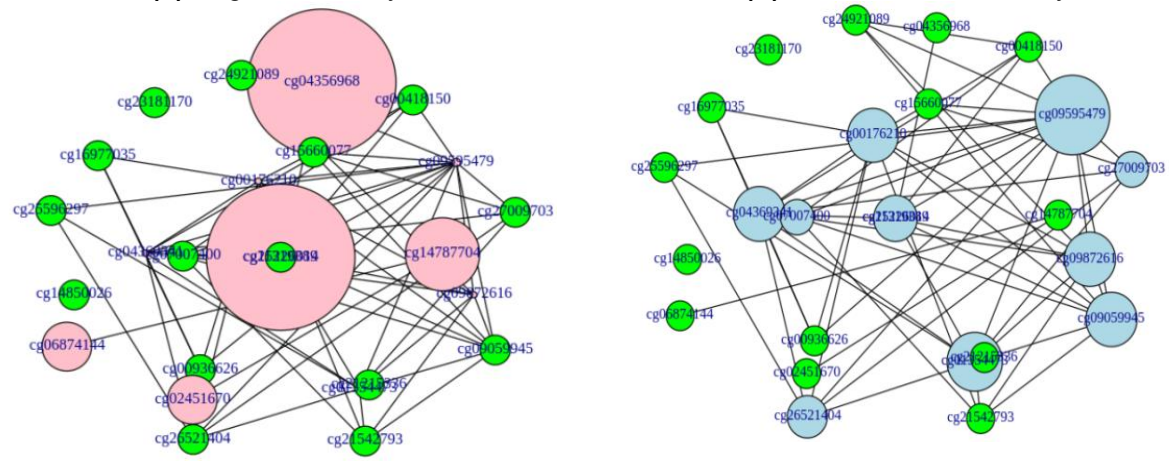


Figure 4. Gene Co-expression network consisting of 25 nodes (genes with gene ids) constructed as per methodology (Reproduced from our previous publication [1] and included to provide biological validation of the identified biomarkers).



(a) Degree Centrality

(b) Betweenness Centrality



(c) Closeness Centrality

(d) Eigenvector Centrality

Figure 5. Centrality based analysis of the GCN **(a)** Red colour nodes: High node weighted degree centrality measure **(b)** Yellow colour nodes: High node weighted betweenness centrality measure **(c)** Pink colour nodes: High node weighted closeness centrality measure **(d)** Blue colour nodes: High node weighted eigenvector centrality measure.

4. Results and Discussion

Proposed node-weighted centrality was used to analyze the gene co-expression network built from the 156 survival-associated genes in the TCGA GBM DNA methylation dataset. The framework combines survival information from the biology and standard graph theoretical centrality measures to prioritize biomarkers relevant to biology. In contrast to the previous study that only considered the topology of the network, the present framework includes biological weights specific to the nodes and combines topological importance with clinical relevance.

Ten biomarkers (APS, MARK2, SEMA3B, SFRP1, SLC44A2, PRLH, DIP2C, PITPNM2, KRTAP4.2 and CTSZ) were consistently identified as priorities across the node-weighted Degree, Closeness, Betweenness, and Eigenvector centrality measures. These biomarkers were found to be the most central in GBM gene co-expression network, which suggests that they are in the top positions both in terms of network topology and in terms of the biological information obtained from survival.

The goal of the proposed node-weighted framework is not to find a new set of biomarkers, but to enhance the prioritization of biomarkers by incorporating biological significance into the calculation of the importance of nodes. Thus, the identification of the same 10 biomarkers as in our previous study, is not a methodological duplication as it indicates the robustness of these candidate biomarkers.

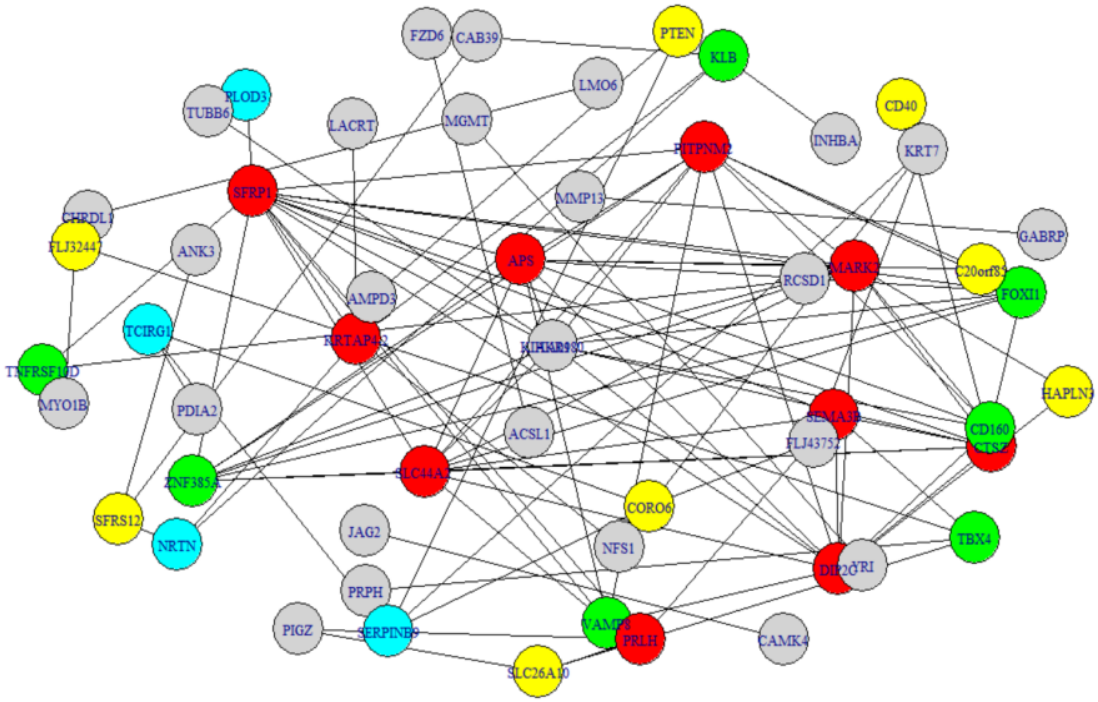


Figure 6. Centrality-measured gene co-expression network analysis: Green color nodes satisfies three centrality criteria, Cyan color nodes satisfies two centrality criteria, yellow color nodes satisfies single centrality criteria and Red nodes meet all four centrality requirements, indicating the most significant DNA Methylation Dataset genes (Reproduced from our previous publication [1] and included to provide biological validation of the identified biomarkers).

Table 3. Univariable examination of the main genes linked to the development of GBM extracted from Methylation Dataset (Reproduced from [1] to provide biological interpretation of the identified biomarkers for the proposed node-weighted centrality framework).

Probe Id.	Gene	HR [95% LCL,95%UCL]	P-Value
cg08097657	SEMA3B	0.016 [0.001 , 0.268]	0.004
cg21663431	SLC44A2	6.847 [1.769 , 25.81]	0.004
cg17998964	MARK2	0.120 [0.023 , 0.517]	0.005
cg06166767	SFRP1	0.121 [0.026 , 0.479]	0.003
cg00025991	DIP2C	0.011 [0.000 , 0.223]	0.004
cg23205633	KRTAP4.2	7.047 [1.801 , 26.51]	0.005
cg16880396	APS	0.038 [0.005 , 0.349]	0.003
cg08176694	PITPNM2	0.050 [0.007 , 0.339]	0.002
cg26427308	PRLH	0.001 [0.000 , 0.167]	0.007
cg23679724	CTSZ	0.008 [0.000 , 0.202]	0.004

The framework, as proposed, is the extension of the conventional network analysis; it proposes a biologically informed ranking strategy that links structural importance to clinical relevance. The biological relevance and the prognostic significance of the prioritized biomarkers are discussed below, and then compared to the previously published conventional centrality-based approach.

After identifying top 20 nodes in each gene co-expression network, nodes shared by all networks that satisfy all centrality measures are pruned from the network. These nodes must meet each criterion of centrality individually. Therefore, the most relevant node, as a genetic biomarker associated with GBM development, was identified as the nodes that represent genes: SEMA3B, MARK2, SLC44A2, APS, PITPNM2, PRLH, CTSZ, SFRP1, DIP2C and KRTAP4.2. This was done by using the nodes to represent these genes. These nodes have been highlighted in red as shown in Figure 6.

Covariate analysis is used in R to determine the impact of these ten significant genes on the onset of disease. The hazard ratios and p-values presented in Table 3 are reproduced from our previous publication [1] and are included here to support the biological interpretation of the proposed node-weighted centrality framework.

Hazard Ratios (HRs) greater than 1 for SLC44A2 and KRTAP4.2 indicate that higher expression levels are associated with poorer overall survival. Lower expression levels of SEMA3B, MARK2, PITPNM2, SFRP1, APS, PRLH, CTSZ and DIP2C are linked to the reduced survival, as indicated by their Hazard Ratios < 1. The goal of the present work was to create and test a node-weighted centrality framework and not to conduct a new survival analysis; therefore, the univariate Cox proportional hazards results from our previous study were deliberately reused as validated biological inputs for the present work. The novelty of the current work is the incorporation of these survival-derived parameters into the network centrality measures for the biologically informed biomarker prioritization

4.1 Survival Analysis

The Kaplan Meier survival analysis is used to compare the overall survival of the key genes. Figure 7. (i,j) shows that K-M curves of genes SLC44A2 and KRTAP4.2 indicate that higher expression levels are associated poorer overall survival of GBM patients, and lower expression levels of genes SEMA3B, MARK2, APS, PITPNM2, PRLH, DIP2C, SFRP1 and CTSZ are associated with poorer overall survival of GBM patients [1].

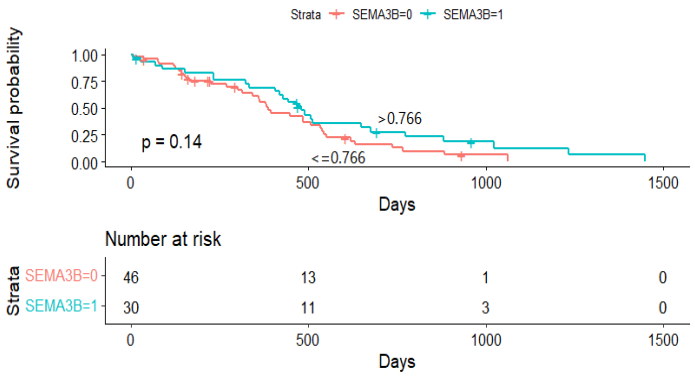
The Kaplan–Meier survival analyses, as well as the cut-off point value of the expression derived from CART, that divided patients into the high and low expression groups, were previously conducted and reported in our previous publication [1]. No recalculation of these analyses was done in the present study. These are presented here for the purposes of biological interpretation and validation of the biomarkers prioritized with the use of the proposed node-weighted centrality framework.

To illustrate the prognostic value of the biomarkers identified by the proposed node-weighted centrality approach, the Kaplan–Meier survival curves presented above are reproduced from our previous study [1]. These curves are for biological interpretation, and are not new survival analysis curves.

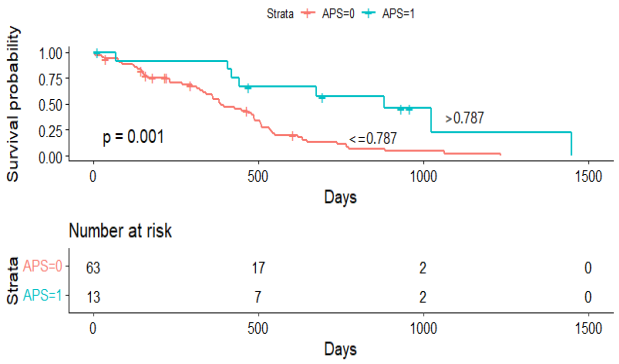
In summary, the biomarkers identified are involved in various biological processes, such as tumor suppression, cell proliferation, remodeling of ECM, signal transduction and epigenetic regulation. Several are known to be involved in glioblastoma or other malignancies like SEMA3B, SFRP1, MARK2 and CTSZ, while others have relatively limited evidence in the literature, such as APS, DIP2C, PRLH, PITPNM2 and KRTAP4.2 [18, 3]. The genes should therefore be seen as exploratory biomarkers and should be further experimentally validated for clinical use.

4.2 Comparison with Previous Study and Advantages of the Node-Weighted Centrality Framework:

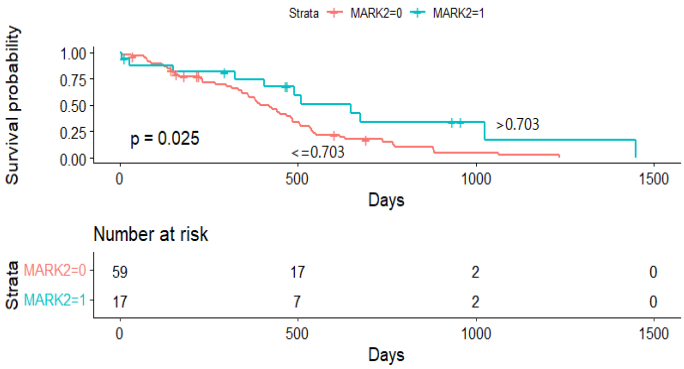
The study presented here shares the same dataset of TCGA GBM methylation data and the same pre-processing protocol as our previous study, but the scientific questions differ significantly between the two studies. The previous studies were dedicated to discovering hub genes by traditional graph theory centrality measures and network topology analysis based on GCN. The present study, however, proposes a novel node-weighted centrality measure incorporating biological information from survival analysis in the network analysis. The same pre-processing pipeline and survival-associated gene selection were kept intentionally, the same 156 genes associated with survival were used for subsequent network construction. Further, the same 10 candidate biomarkers (SEMA3B, SLC44A2, APS, MARK2, SFRP1, PITPNM2, PRLH, CTSZ, DIP2C and KRTAP4.2) were found in the proposed framework, which shows the robustness and stability of these biomarkers between different centrality formulations. The proposed methodology is not intended to replace the biomarkers already identified, but is intended to offer an alternate biologically-informed ranking approach, combining structural network and survival significance. The novelty of the present study is thus not in the discovery of a novel set of biomarkers, but rather in the establishment of a novel node-weighted centrality framework that is optimized for survival, incorporating a biological component in node prioritization.



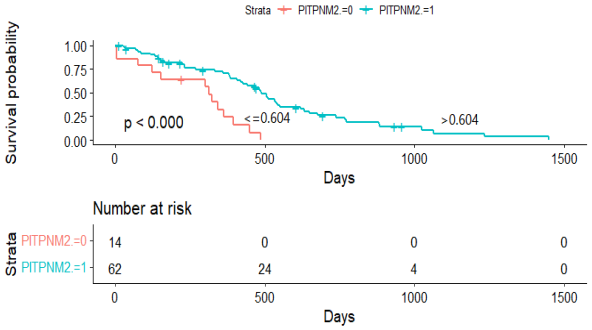
(a) SEMA3B



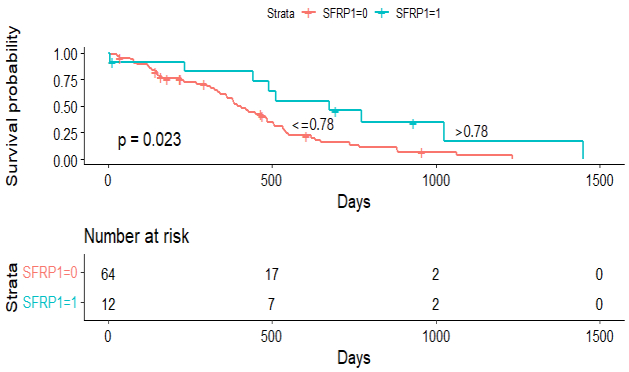
(b) APS



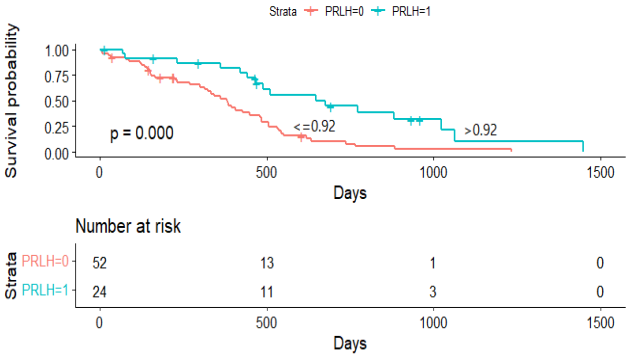
(c) MARK2



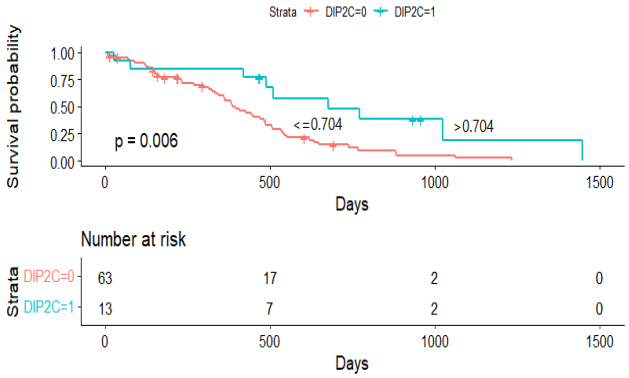
(d) PITPNM2



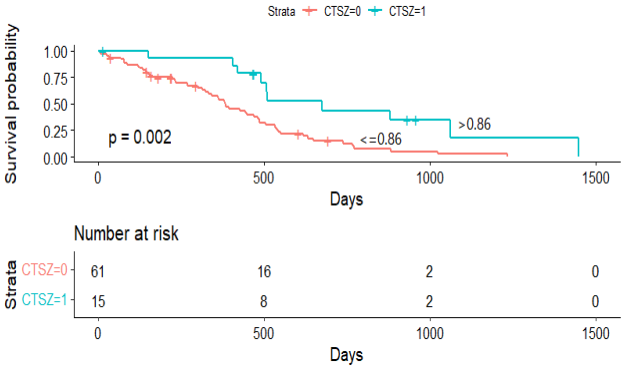
(e) SFRP1



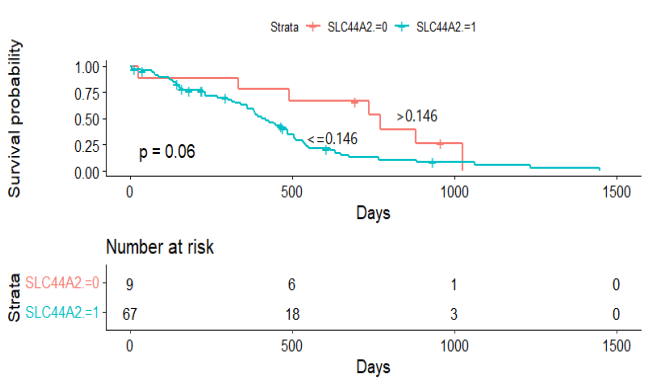
(f) PRLH



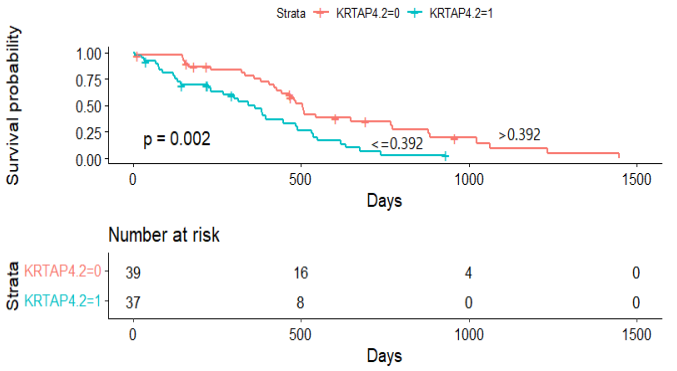
(g) DIP2C



(h) CTSZ



(i) SLC44A2



(j) SLC44A2

Figure 7. Comparing the overall survival of key genes at their threshold levels using Kaplan-Meier curves from the DNA Methylation Dataset (Reproduced from [1] to provide biological interpretation of the identified biomarkers for the proposed node-weighted centrality framework).

In the previous study only conventional graph-theoretic centrality measures were used for assessing the importance of the nodes. The situation is different in the current framework, however, where the p-values are derived from survival are used to weight the nodes when calculating centrality. The same set of biomarkers was detected, but the proposed formulation is a variant of topology-based network analysis, with a biologically motivated formulation that incorporates the significance of the biomarkers with respect to the survival.

4.3 Study Limitations and Future Scope

There are some limitations to this study which should be taken into account when considering the findings. Firstly, the suggested node-weighted centrality framework was validated in one TCGA GBM DNA methylation cohort. TCGA is a legitimate public repository of cancer genome data but the proposed methodology should be validated in independent GBM cohorts to further support the generalizability of the proposed methodology. The study intentionally reused the previously validated preprocessing pipeline, the previously validated gene survival selection and the previously validated survival statistics to ensure methodological consistency and to assess the proposed node-weighted centrality framework on a common biological foundation. The focus of the present work is thus methodological, not on the discovery of a new set of biomarkers.

In this study, the gene co-expression network was built with absolute Pearson correlation cut-off value of 0.5. In the present study, we did not explore the effect of different threshold values on network topology or biomarker prioritization, but this choice has been made in gene co-expression studies to keep biologically interesting gene associations. In addition, the proposed node-weighted framework includes biological information derived from univariate Cox regression, which is generated through the survival process.

Lastly, this study is computational. For the prioritized biomarkers, their biological roles and value for prognostication would need to be experimentally verified and further studies on prospective clinical applications should be undertaken. The proposed framework will be validated using independent patient cohorts in future research studies, which would involve the inclusion of other clinical and multi-omics data sets, the use of other network construction approaches and weighting systems, and the application of the proposed node-weighted centrality approach to additional cancer types. These studies will also help to further validate biologically informed network analysis for biomarker prioritization, its robustness, generalizability, and applicability.

5. Conclusion

This study proposed a biologically-informed node-weighted centrality framework for biomarker prioritization using survival-derived biological information and incorporated such information in the topology of gene co-expression networks. This proposed approach builds on our previously published framework for centrality-based biomarker discovery pipeline by adding the node-specific survival information to the computation of graph-theoretic centrality measures and preserves the well-established biological underpinning. The framework consistently identified as top candidates 10 candidate biomarkers, SEMA3B, SLC44A2, APS, MARK2, SFRP1, PITPNM2, PRLH, CTSZ, KRTAP4.2 and DIP2C which were both topologically important and biologically relevant in the GBM TCGA DNA methylation dataset.

The discovery of the same set of biomarkers in both traditional and node-weighted centrality measures indicates the general strength and consistency of these potential biomarkers. It does not propose to identify a new set of biomarkers, however, but a biologically informed approach to prioritize candidate genes by

integrating survival-associated biological information with network topology.

This study is aimed to make a methodological contribution to network-based biomarker prioritization. While the identified biomarkers are clinically relevant by being previously validated with survival analysis, further validation with independent patient groups, multivariable survival analysis and experimental studies will be required before their clinical translation. In future studies, we will explore other node-weighting transformations, alternative network construction approaches, and incorporation of multi-omics and clinical data to further enhance the robustness and applicability of biologically informed network analysis to discover biomarkers in GBM and other cancers.

References

[1] S.S. Redekar, S.L. Varma, A. Bhattacharjee, Gene co-expression network construction and analysis for identification of genetic biomarkers associated with glioblastoma multiforme using topological findings. Journal of the Egyptian National Cancer Institute, 35(1), (2023) 22. https://doi.org/10.1186/s43046-023-00181-4

[2] K. Cullison, (2024). Machine Learning and Radiomics Investigations of Daily MRI-Guided Glioblastoma Radiotherapy. PhD dissertation, University of Miami.

[3] A.C. Tan, D.M. Ashley, G.Y. López, M. Malinzak, H.S. Friedman, M. Khasraw, Management of glioblastoma: State of the art and future directions. CA: A Cancer Journal for Clinicians, 70(4), (2020) 299–312. https://doi.org/10.3322/caac.21613

[4] R.G. Dumitrescu, Epigenetic Markers of Early Tumor Development. In: Dumitrescu, R., Verma, M. (eds) Cancer Epigenetics. Methods in Molecular Biology, Humana Press, Totowa, NJ, 863. (2012) 3–14. https://doi.org/10.1007/978-1-61779-612-8_1

[5] B.K. Ahir, H.H. Engelhard, S.S. Lakka, Tumor development and angiogenesis in adult brain tumor: glioblastoma. Molecular Neurobiology, 57(5), (2020) 2461–2478. https://doi.org/10.1007/s12035-020-01892-8

[6] I. Mani, V. Singh, Applications of bioinformatics in epigenetics. Progress in Molecular Biology and Translational Science, 198, (2023) 1–13. https://doi.org/10.1016/bs.pmbts.2023.03.023

[7] X. Yu, H. Zhao, R. Wang, Y. Chen, X. Ouyang, W. Li, Y. Sun, A. Peng, Cancer epigenetics: From laboratory studies and clinical trials to precision medicine. Cell Death Discovery, 10(1), (2024) 28. https://doi.org/10.1038/s41420-024-01803-z

[8] M.A. Hayat, (2011). Tumors of the Central Nervous System, Volume 1: Gliomas: Glioblastoma (Part 1). Springer Science & Business Media.

[9] A. Rusak, B. Wiatrak, K. Krawczyńska, T. Górnicki, K. Zagórski, Ł. Zadka, W. Fortuna, Starting points for the development of new targeted therapies for glioblastoma multiforme. Translational Oncology, 51, (2025) 102187. https://doi.org/10.1016/j.tranon.2024.102187

[10] O. Mason, M. Verwoerd, Graph theory and networks in biology. IET Systems Biology, 1(2), (2007) 89–119. https://doi.org/10.1049/iet-syb:20060038

[11] S. Ranganathan, K. Nakai, C. Schonbach, (2018). Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics. Elsevier.

[12] G. Muzio, L. O'Bray, K. Borgwardt, Biological network analysis with deep learning. Briefings in Bioinformatics, 22(2), (2021) 1515–1530. https://doi.org/10.1093/bib/bbaa257

[13] H.A. Chowdhury, D.K. Bhattacharyya, J.K. Kalita, (Differential) co-expression analysis of gene expression: a survey of best practices. IEEE/ACM Transactions on Computational Biology and Bioinformatics, IEEE, 17(4), (2019) 1154–1173. https://doi.org/10.1109/TCBB.2019.2893170

[14] G.G. Lemoine, M.P. Scott-Boyer, B. Ambroise, O. Périn, A. Droit, GWENA: gene co-expression networks analysis and extended modules characterization in a single Bioconductor package. BMC Bioinformatics, 22(1), (2021) 267. https://doi.org/10.1186/s12859-021-04179-4

[15] T.P. de Souza, G.S. de Araújo, L. Magalhães, G.C. Cavalcante, A. Ribeiro-dos-Santos, C. Sena-dos-Santos, C.S. Silva, G.L. Eufraseo, A. de Freitas Escudeiro, G.B. Soares-Souza, B.L. Santos-Lobato, Unveiling differential gene co-expression networks and its effects on levodopa-induced dyskinesia. iScience, 27(9), (2024) 110835. https://doi.org/10.1016/j.isci.2024.110835

[16] S.K. Ray, (2009). Glioblastoma: Molecular Mechanisms of Pathogenesis and Current Therapeutic Strategies. Springer. https://doi.org/10.1007/978-1-4419-0410-2

[17] C.F.A. Negre, U.N. Morzan, H.P. Hendrickson, R. Pal, G.P. Lisi, J.P. Loria, I. Rivalta, J. Ho, V.S. Batista, Eigenvector centrality for characterization of protein allosteric pathways. Proceedings of the National Academy of Sciences, 115(52), (2018) E12201–E12208. https://doi.org/10.1073/pnas.1810452115

[18] D.N. Louis, A. Perry, P. Wesseling, D.J. Brat, I.A. Cree, D. Figarella-Branger, C. Hawkins, H.K. Ng, S.M. Pfister, G. Reifenberger, R. Soffietti, The 2021 WHO classification of tumors of the central nervous system: a summary. Neuro-Oncology,

23(8), (2021) 1231–1251. https://doi.org/10.1093/neuonc/noab106

[19] S. Shah, A. Nag, S.V. Sachithanandam, B. Lucke-Wold, Predictive and prognostic significance of molecular biomarkers in glioblastoma. Biomedicines, 12(12), (2024) 2664. https://doi.org/10.3390/biomedicines12122664

[20] D. Gibson, A.H. Vo, H. Lambing, P. Bhattacharya, P. Tahir, F.F. Chehab, N. Butowski, A systematic review of high impact CpG sites and regions for MGMT methylation in glioblastoma [A systematic review of MGMT methylation in GBM]. BMC Neurology, 24(1), (2024) 103. https://doi.org/10.1186/s12883-024-03605-3

[21] Cancer Genome Atlas Research Network, R. McLendon, A. Friedman, D. Bigner, E.G. Van Meir, D.J. Brat, G. Mastrogianakis, J.J. Olson, T. Mikkelsen, N. Lehman, K. Aldape, W.K. Yung, O. Bogler, S. VandenBerg, M. Berger, M. Prados, P.W. Laird, L. Cope, J.G. Herman, D.J. Weisenberger, F. Pan, D. Van Den Berg, L. Van Neste, Y.J. Mi, K.E. Schuebel, S.B. Baylin, Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature, 455(7216), (2008) 1061–1068. https://doi.org/10.1038/nature07385

[22] M.A. Jacome, Q. Wu, Y. Piña, A.B. Etame, Evolution of molecular biomarkers and precision molecular therapeutic strategies in glioblastoma. Cancers, 16(21), (2024) 3635. https://doi.org/10.3390/cancers16213635

[23] P.H. Model, (2021). Cox proportional hazards model.

[24] S. Van Dam, U. Vosa, A. Van der Graaf, L. Franke, J.P. de Magalhaes, Gene co-expression analysis for functional classification and gene–disease predictions. Briefings in Bioinformatics, 19(4), (2018) 575–592. https://doi.org/10.1093/bib/bbw139

[25] S. Feng, L. Zhu, J. Gu, K. Kou, Y. Liu, G. Zhang, H. Lu, H. Zhang, R. Sun, Exploration of M2 macrophage-related biomarkers and a candidate drug for glioblastoma using high-dimensional weighted gene co-expression network analysis. Frontiers in Pharmacology, 16, (2025) 1587258. https://doi.org/10.3389/fphar.2025.1587258

[26] J. Luan, D. Zhang, B. Liu, A. Yang, K. Lv, P. Hu, H. Yu, A. Shmuel, C. Zhang, G. Ma, Exploring the prognostic value and biological pathways of transcriptomics and radiomics patterns in glioblastoma multiforme. Heliyon, 10(13), (2024) e33760. https://doi.org/10.1016/j.heliyon.2024.e33760

[27] Z. Wu, X. Wang, H. Wu, S. Du, Z. Wang, S. Xie, R. Zhang, G. Chen, H. Chen, Identification of CREB5 as a prognostic and immunotherapeutic biomarker in glioma through multi-omics pan-cancer analysis. Computers in Biology and Medicine, 173, (2024) 108307. https://doi.org/10.1016/j.compbiomed.2024.108307

[28] F. Bu, J. Zhong, R. Guan, Biomarkers in glioblastoma and degenerative CNS diseases: defining new advances in clinical usefulness and therapeutic molecular target. Frontiers in Molecular Biosciences, 12, (2025) 1506961. https://doi.org/10.3389/fmolb.2025.1506961

[29] D. Koschützki, F. Schreiber, Centrality analysis methods for biological networks and their application to gene regulatory networks. Gene Regulation and Systems Biology, 2, (2008) GRSB-S702. https://doi.org/10.4137/GRSB.S702

[30] Y. Wang, Z. Wang, X. Yu, X. Wang, J. Song, D.J. Yu, F. Ge, MORE: a multi-omics data-driven hypergraph integration network for biomedical data classification and biomarker identification. Briefings in Bioinformatics, 26(1), (2025) bbae658. https://doi.org/10.1093/bib/bbae658

[31] M.N. Zhang, M.J. Xue, B.Z. Zhou, J. Xu, H.K. Sun, J.H. Wang, Y.Y. Wang, Comprehensive review of Bayesian network applications in gastrointestinal cancers. World Journal of Clinical Oncology, 16(6), (2025) 104299. https://doi.org/10.5306/wjco.v16.i6.104299

[32] M.E.J. Newman, (2010). Networks: An Introduction. Oxford University Press. https://doi.org/10.1093/oso/9780198805090.001.0001

[33] J. Hu, J. Yang, N. Hu, Z. Shi, T. Hu, B. Mi, H. Wang, W. Chen, Identification and verification of key genes associated with temozolomide resistance in glioblastoma based on comprehensive bioinformatics analysis. Iranian Journal of Biotechnology, 22(4), (2024) e3892. https://doi.org/10.30498/ijb.2024.448826.3892

[34] P. Zohari, M.H. Chehreghani, (2025). Graph Neural Networks in Multi-Omics Cancer Research: A Structured Survey. arXiv preprint arXiv:2506.17234. https://doi.org/10.48550/arXiv.2506.17234

[35] Z. Lan, X. Li, X. Zhang, Glioblastoma: An update in pathology, molecular mechanisms and biomarkers. International Journal of Molecular Sciences, 25(5), (2024) 3040. https://doi.org/10.3390/ijms25053040

[36] H. Liu, D. Zhou, Y. Ou, S. Chen, Y. Long, T. Yuan, Y. Li, Y. Tan, Multiple signaling pathways in the frontiers of lung cancer progression. Frontiers in Immunology, 16, (2025) 1593793. https://doi.org/10.3389/fimmu.2025.1593793

[37] X. Tian, Z. Li, Y. Tian, Z. Cui, Z. Wang, C. Peng, Celebrating 25 years of Briefings in Bioinformatics: a comprehensive bibliometric analysis. Briefings in Bioinformatics, 27(2),

(2026) bbag117. https://doi.org/10.1093/bib/bbag117

[38] A. Ma, X. Wang, J. Li, C. Wang, T. Xiao, Y. Liu, H. Cheng, J. Wang, Y. Li, Y. Chang, J. Li, D. Wang, Y. Jiang, L. Su, G. Xin, S. Gu, Z. Li, B. Liu, D. Xu, Q. Ma, Single-cell biological network inference using a heterogeneous graph transformer. Nature Communications, 14(1), (2023) 964. https://doi.org/10.1038/s41467-023-36559-0

[39] C. Hutter, J.C. Zenklusen, The Cancer Genome Atlas: Creating Lasting value beyond its Data. Cell, 173(2), (2018) 283–285. https://doi.org/10.1016/j.cell.2018.03.042

[40] R. Swaminathan, H. Brenner, Statistical Methods for Cancer Survival Analysis. IARC Scientific Publications, 162, (2011) 7–13.

[41] D.R. Youlden, P.D. Baade, C.A. McBride, J.D. Pole, A.S. Moore, P.C. Valery, A. Young, J.F. Aitken, Changes in Cancer Incidence and Survival Among Aboriginal and Torres Strait Islander Children in Australia, 1997–2016. Pediatric Blood and Cancer, 69(4), (2022) e29492. https://doi.org/10.1002/pbc.29492

[42] M. Stock, N. Popp, J. Fiorentino, A. Scialdone, Topological Benchmarking of Algorithms to Infer Gene Regulatory Networks from Single-Cell RNA-seq Data. Bioinformatics, 40(5), (2024) btae267. https://doi.org/10.1093/bioinformatics/btae267

[43] M. Saberi, R. Khosrowabadi, A. Khatibi, B. Misic, G. Jafari, Topological Impact of Negative Links on the Stability of Resting-State Brain Network. Scientific Reports, 11(1), (2021) 2176. https://doi.org/10.1038/s41598-021-81767-7

[44] J.P. Klein, Survival Analysis Methods in Cancer Studies. In: Beam, C. (edts). Bio Statistical Applications in Cancer Research. Cancer Treatment and Research, Springer, Boston, MA. 113, (2002) 37–57. https://doi.org/10.1007/978-1-4757-3571-0_3

[45] B.O. Omolo, S.O. Manda, Application of biostatistics and epidemiological methods for cancer research in Sub-Saharan Africa. Frontiers in Public Health, 10, (2022) 1069098. https://doi.org/10.3389/fpubh.2022.1069098

[46] O. Kardos, A. London, T. Vinkó, Stability of network centrality measures: a numerical study. Social Network Analysis and Mining, 10(1), (2020) 1–17. https://doi.org/10.1007/s13278-020-00693-0

[47] J. Kim, M. Hastak, Social network analysis: Characteristics of online social networks after a disaster. International Journal of Information Management, 38(1), (2018) 86–96. https://doi.org/10.1016/j.ijinfomgt.2017.08.003

[48] T. Vetrík, Extended study of biological networks using graph theory. Proyecciones (Antofagasta), 40(5), (2021) 1211–1225.

[49] M.V.C. Santos, A.S. Feltrin, I.C. Costa-Amaral, L.R. Teixeira, J.A. Perini, D.C. Martins Jr, A.L. Larentis, Network analysis of biomarkers associated with occupational exposure to benzene and malathion. International Journal of Molecular Sciences, 24(11), (2023) 9415. https://doi.org/10.3390/ijms24119415

[50] T. Rout, A. Mohapatra, M. Kar, D.K. Muduly, Essential Protein Identification in Cancer: A Graph-Based Approach Integrating Topological and Biological Features in PPI Networks. SN Computer Science, 5(7), (2024) 947. https://doi.org/10.1007/s42979-024-03312-3

[51] J. Jiang, X. Sun, W. Wu, L. Li, H. Wu, L. Zhang, G. Yu, Y. Li, Construction and Application of a Co-Expression Network in Mycobacterium Tuberculosis. Scientific Reports, 6(1), (2016) 28422. https://doi.org/10.1038/srep28422

[52] B. Zhang, S. Horvath, A General Framework for Weighted Gene Co-Expression Network Analysis. Statistical Applications in Genetics and Molecular Biology, 4(1), (2005). https://doi.org/10.2202/1544-6115.1128

Authors Contribution Statement

Seema Sandeep Redekar: Conceptualization, Methodology, Software, Data Curation, Writing - Original Draft. Smita Dange: Visualization, Investigation. Madhura Vyawahare: Software, Validation, Writing - Review & Editing. All authors have read and agreed to the published version of the manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes

About the License

© The Author(s) 2026. The text of this article is open access and licensed under a Creative Commons Attribution 4.0 International License.