A new method for patient stratification based on multi-layer network modeling and molecular data integration

Manuela Petti[a], Davide Mascolo[b], Caterina Alfano[c], Lorenzo Farina[a]

[a]*Dept. of Computer and Management Engineering, Sapienza University of Rome, Italy.*
[b]*Ph.D. Program in Data Science, Sapienza University of Rome, Italy.*
[c]*Ph.D. Program in Network Oncology and Precision Medicine, Sapienza University of Rome, Italy.*

Abstract

With the development of increasingly complex and heterogeneous diseases, there has been a need to use computational methods to support traditional medicine in choosing optimal treatment for patients. This work focused on developing a new approach based on multi-layer network modeling and integrating molecular data for patient stratification. We applied the proposed method to combine mRNA expression, DNA methylation, and microRNA (miRNA) expression data for six cancers. The proposed method outperformed the state-of-the-art algorithm Similarity Network Fusion (SNF) obtaining a better stratification of Glioblastoma Multiforme (GBM) patients for overall survival, and it ensures meaningful stratification of patients with enrichment of clinical variables also for Breast Invasive Carcinoma (BIC), Kidney Renal Clear Cell Carcinoma (KRCCC), Lung Squamous Cell Carcinoma (LSCC), and Colon Adenocarcinoma (COAD). Considering the independent dataset composed of Non-Small Cell Lung Cancer (NSCLC) patients, there is a significant stratification of patients ($p-value \leq 0.05$) with clinical characterization by sex, and age that opens up future debates and studies on the topic of sexual dimorphism. The proposed method successfully addresses both challenges of patient stratification and biomedical data integration with potential clinical impact in therapy identification.

Keywords: Networks, Omics, Patient Stratification, Cancer, Immunotherapy

1. Introduction

Precision medicine is considered an emerging approach for disease treatment that considers individual characteristics of patients in terms of molecular profile, lifestyle, and clinical history (1) (2). The advantage of this approach is the use of computational methods that can integrate several data sources to create a comprehensive view of the disease. The emerging paradigm driving these methods is network science, which models patient relationships across multiple data layers using network approaches to achieve meaningful patient stratification (3). This is an important goal for biomedical research, given that within a unique cohort of patients affected by the same disease, it is possible to identify subgroups based on their clinical features and/or molecular profiles, especially for complex and heterogeneous diseases such as cancer. Patient Similarity Network (PSN) (8) is the most common paradigm used for patient stratification where each node in a network is seen as a patient and the edges between nodes represent the similarity between biological samples. As reported in Fig. 1, this approach starts with several data layers for which the idea is to build a PSN. The next step is integrating the different PSNs into a single multi-layer network model that will be the patient stratification step's input.

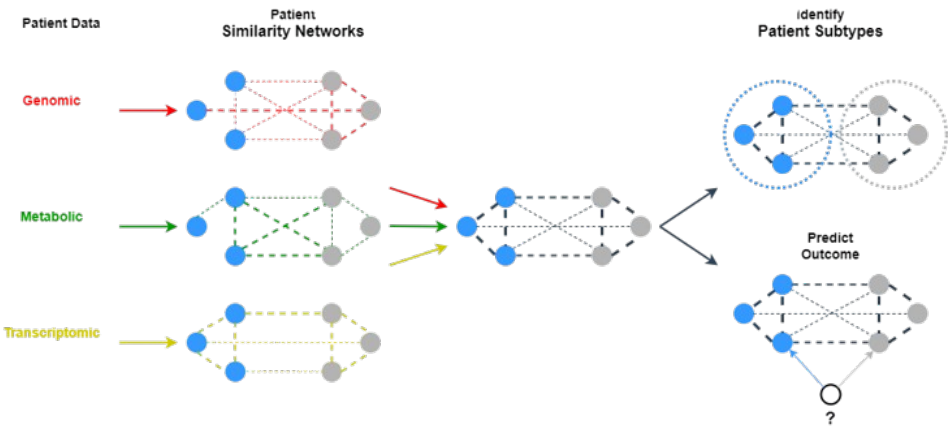


Figure 1: Patient Similarity Network (PSN) framework

Before this study, several PSN-based methods were used for patient stratification, with the SNF algorithm representing the state-of-the-art (4). Wang and colleagues developed it to integrate a set of unimodal PSNs and it overcomes the limitations of Network-Based Stratification (5) related to the use of a not well-defined network to map the mutations. The SNF has been applied in many different applications (6) (7) highlighting the potential and at the same time the limitations of this method, and several studies proposed some changes to improve and expand SNF. Building on this knowledge gap, the authors propose a new approach for patient stratification that leverages multi-layer network modeling and molecular data integration to surpass state-of-the-art method. The proposed conceptual framework, not only exceeds SNF method but turns out to be a replicable and extendable approach to other datasets having the same characteristics, to achieve meaningful stratification for clinical variables, of the patients analyzed.

2. Materials and Methods

2.1. Data

To test the proposed method, the authors have used two independent cohorts of patients suffering from different types of cancer. We downloaded the first cohort of patients, used by Wang and colleagues (4), from the TCGA website selecting the three data types: *gene expression*, *miRNA expression*, and *DNA methylation* for five different cancers.

Cancer Type	Patients Number	mRNA Expression	DNA Methylation	miRNA Expression
GBM	215	12.042	1.491	534
BIC	105	17.814	23.094	1.046
KRCCC	124	20.532	24.976	1.046
LSCC	105	12.042	27.578	1.046
COAD	92	17.814	27.578	705

Table 1: Patient number used in our experiments based on (4)

To validate the proposed method, in addition to the data used in (4), it was used an independent dataset containing 393 molecular data for NSCLC patients and already analyzed by the authors in the previous work (9).

Cancer Type	Patients Number	Rna-Seq	WES	Curated Signatures	Myeloid Signatures	Immune Signatures
NSCLC	393	57523	52	7	18	11

Table 2: Patient number used in our experiments based on (10)

The set of GBM patients is composed mostly of men, precisely 62% of men and women represented by 38% as reported in Fig. 2. For NSCLC data, the clinical response patients' information ranged in age from 29 to 90 years, composed of 182 men and 207 women (Fig. 3). All the patients were treated with an anti-PD-(L)1 agent for immunotherapy considering the genomic and transcriptomic profile.

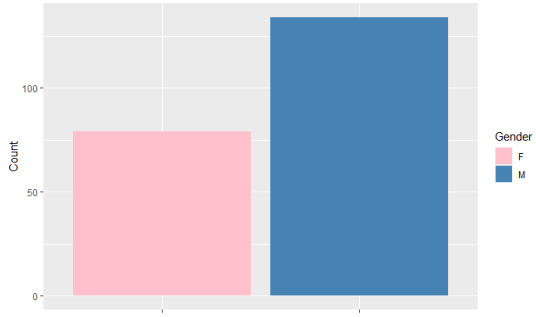


Figure 2: Gender Distribution in GBM Patients

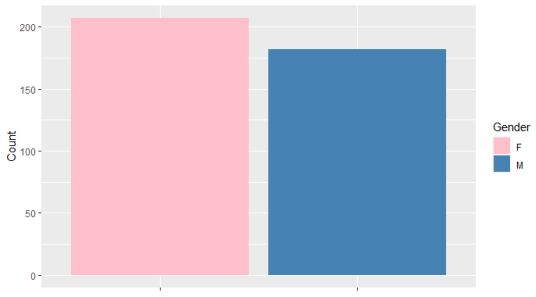


Figure 3: Gender Distribution in NSCLC Patients

2.2. Framework

For each data type in the two datasets, we generated the PSN by calculating patient similarity using Euclidean distance. Then we stack the unimodal PSNs obtaining a multi-layer network model and we applied for 1000 iterations Generalized Louvain algorithm (11) to identify a common partition. The output of (11) is a *patient-by-iteration* matrix, which is processed using the Lloyd k-means algorithm to identify the most stable partition across 1000 iterations and assess the survival curves of the resulting clusters. The Generalized Louvain algorithm was applied using the MATLAB package developed by Lucas G. and colleagues (12). Using this approach, similarity matrices for each data layer ($patient-patient$) must be constructed. These matrices are then utilized to generate the multi-layer model. To accomplish this, two parameters need to be defined: the intralayer parameter γ and the interlayer parameter ω. In this study, γ is set to 0.82, while ω equals 1. The output of the generalized Louvain algorithm is the partition matrix represented in Fig. 4, where each row represents a different patient, and the columns represent the number of iterations. Considering Fig. 4, it is clear that many patients are classified into the same cluster, index of homogeneity of the result. Moreover, in most cases, the algorithm identified two or three clusters, reaching a maximum of five.

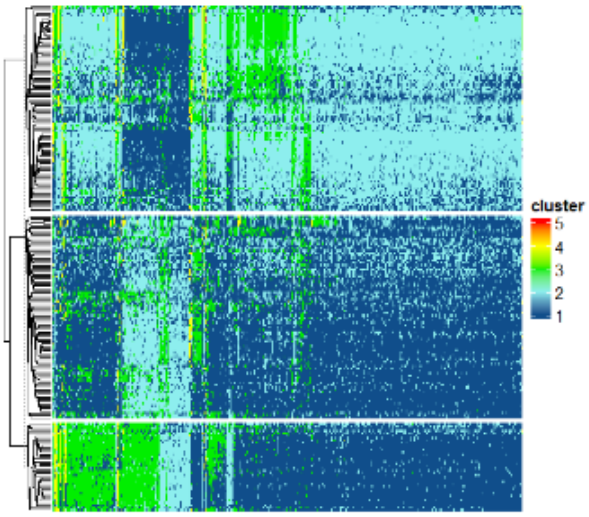


Figure 4: Heatmap Patient for iteration

3. Results

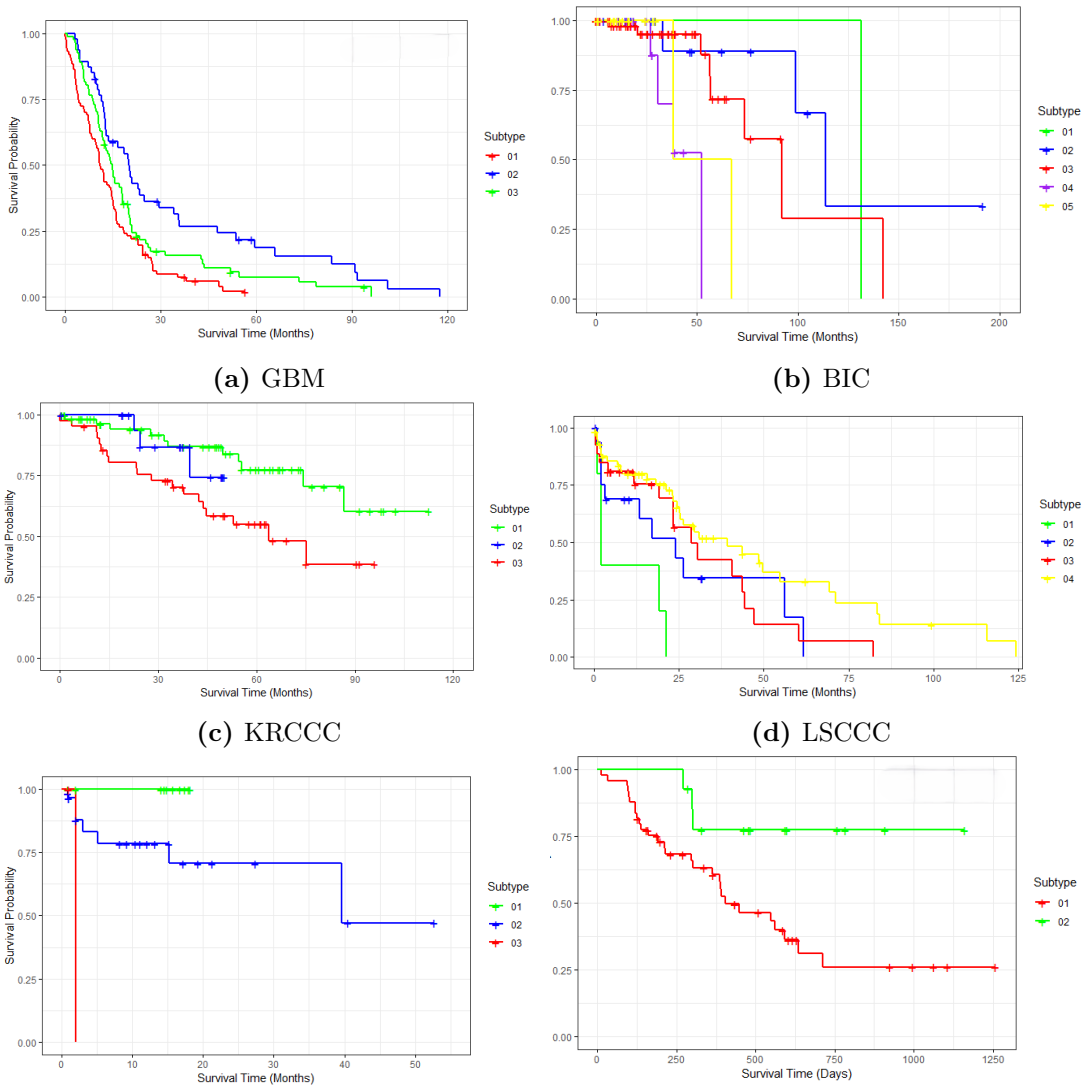
3.1. Survival Analysis

To compare the method proposed in this work and the SNF methodology, we used the same evaluation method of (4), the $p - value$ associated with the Cox log-rank test. This comparison demonstrates both that data integration is statistically significant and necessary to achieve the desired stratification and the validity of the proposed method compared with SNF.

Cancer Type	Clusters Number	SNF	Proposed Method
GBM	3	$2.0x10^{-4}$	$4.0x10^{-5}$
BIC	5	$1.1x10^{-3}$	$3.4x10^{-3}$
KRCCC	3	$2.9x10^{-2}$	$2.8x10^{-2}$
LSCC	4	$2.0x10^{-2}$	$8.3x10^{-4}$
COAD	3	$8.8x10^{-4}$	$2.7x10^{-2}$

Table 3: Analysis using Cox long-rank test P-values based on (4)

Before proceeding with a qualitative analysis of the results to compare the proposed method with the SNF, the authors ensured that the proposed method had statistically significant performance, from a quantitative point of view, with the reference method. At this stage, the study aims to validate the proposed method from a qualitative view by comparing the survival curves obtained for different diseases. Considering the GBM patients, Fig. 5.a shows the overall survival of the 3 subgroups of patients identified with the proposed method. No alternative combination of the three data layers (use of single datasets or pairs) has returned a stratification associated with such a clear separation of the survival curves. Also, SNF returned a significant stratification, but two subtypes are not well separated (4). Regarding the survival curves of the other diseases analyzed, the profiles identified in KRCCC (5.c), LSCC (5.d), and NSCLC (5.f) deserve special attention. In particular, for KRCCC (5.c), the distinction between the 3 clusters identified by our method appears to be fairly well separated from the early months of survival, with only a slight overlap around month 30 between subtype one and subtype two. The SNF method, on the other hand, on the same pathology does not distinguish well between the curves identified as subtype two and subtype three but overlaps them until the 30th month of survival, the point where separation begins. Regarding LSCC (5.d), on the other hand, our method identifies the four clusters of the baseline study, but with some differences. It is interesting what happens in cluster one identified, which has a probability of survival going to zero very quickly, compared with the cluster with lower survival identified by the SNF method, which also blurs with subtype four by overlapping in several places. In addition, subtype four in our study exhibits a survival pattern that is quite different and distinct from the other subtypes, especially regarding survival time, although there is slight contact with subtype two and subtype three at some points. Considering the independent NSCLC (5.f) cohort and compared with the survival curves obtained with the SNF method, there is a clear separation between the two groups identified from the beginning, and the two clusters never overlap, an effect that was the case with the SNF method.



(a) GBM **(b)** BIC

(c) KRCCC **(d)** LSCCC

(e) COAD **(f)** NSCLC

Figure 5: Kaplan-Meier survival curves (p-values reported in Table 4)

3.2. Clinical Enrichment Analysis

After consolidating the proposed methodology also from a qualitative point of view, looking at the survival curves, and improving in some pathologies the separation obtained with the SNF method, the objective is identified through a clinical enrichment which are the most relevant factors for the stratifications obtained. Enrichment for clinical variables is done by considering the subtypes identified for each disease analyzed and the clinical variables associated with them. Specifically, once the patient IDs within each subtype have been identified, the corresponding clinical profile is extracted, and characterization by discriminant variables is performed.

3.2.1. GBM Patients

For patients with GBM (Glioblastoma Multiforme), the discriminant variable is *Age*. Looking at Fig. 6, subtype two is the youngest (*median age = 41 years*) but with the greatest intra-group variability. Subtype one, which corresponds to the lowest survival probability curve, is older (*median age = 60 years*). In the middle, there is subtype three (*median age = 55.5 years*). This analysis finds confirmation in the survival curves

shown above. In particular, the identified subtype one compound characterized by the highest median age compared with all the three identified groups is also the subtype with the lowest survival probability (Fig. 5.a). In contrast, subtype two, characterized by wide variability and the lowest median age, has the highest survival probability, while in an intermediate situation is subtype three, both in terms of median age and survival probability.

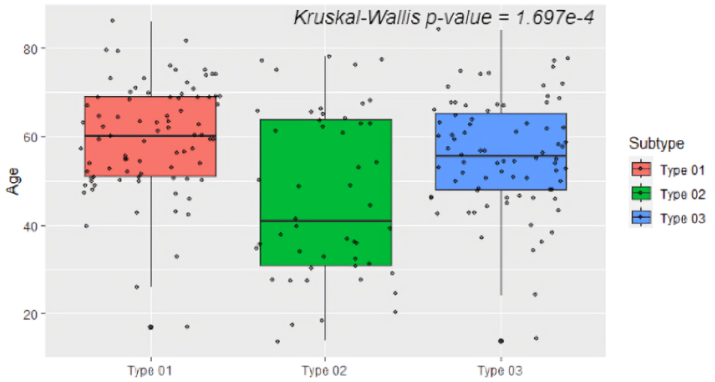


Figure 6: Age Distribution for GBM subtypes

3.2.2. KRCCC Patients

Regarding patients affected by KRCCC (Kidney Renal Clear Cell Carcinoma), enrichment for clinical variables returns several variables as a discriminant for the patients' stratification (Table 4). After testing the statistical significance of the observed effect in these variables, an extraction of the clinical characteristics of the identified subtypes was performed to study their distribution.

Variable	P-value
Stage	$5.0x10^{-2}$
Vital Status	$1.0x10^{-2}$
Synchronous Malignancy	$2.0x10^{-2}$
Pathologic T (AJCC Staging System)	$2.0x10^{-2}$

Table 4: Relevant clinical features for GBM patients stratification

It is possible to see in Table 5, that subtype one, which is characterized by the greatest survival probability is mainly composed of patients affected by stage I of the disease, while subtype three is relevant in patients affected by the disease by stage III and IV. Regarding subtype two, this catches some patients becoming stages I, II, and III, explainable as a noise subtype. This result shows the significance and impact of the disease's stage on survival rate. Especially for patients with early-stage colon cancer, there is a higher survival rate than for other stages, and this underscores the critical importance of early prognosis and therapy management strategies (17). Regarding the variable *Pathologic T* for AJCC staging system, it refers to TNM Classification of Malignant Tumors (TNM), which is a global standard to classify the anatomical extent of the spread of cancers, developed by the American Joint Committee on Cancer (AJCC). The T measurement of this variable describes the size of the original (primary) tumor and any spread of cancer into nearby tissue. Subtype one is characterized by a high concentration of patients (n:18) with T1b staging (tumor $>$ 4 but $\leq$ 7 cm in greatest dimension, limited to the kidney), and T1a (n:14) (tumor $\leq$ 4 cm in greatest dimension, limited to the kidney), which means that the size and the extension of the tumor are low. On the opposite side, in subtype three there is a major concentration of patients (n:14) affected by T3a (tumor extends into the renal vein or its segmental branches, or invades the pelvicalyceal system, or invades perirenal and/or renal sinus fat but not beyond Gerota's fascia) (15). This aspect confirms and explains what we see with survival curves and emphasizes the lowest probability of survival for subtype three. This factor turns out to be of paramount importance, in fact as shown in (18), the risks of malignancy and high-grade tumors increase with tumor size. Whereas, patients with small renal masses have a low risk of harboring a malignancy, which can be helpful when deciding on therapy (16).

Subtype	Stage	Count
01	I	32 (26.2%)
01	III	13 (10.7%)
01	II	9 (7.38%)
01	IV	7 (5.74%)
02	I	13 (10.7%)
02	III	4 (3.28%)
02	II	2 (1.64%)
03	III	16 (13.1%)
03	I	15 (12.3%)
03	II	6 (4.92%)
03	IV	4 (3.28%)
03	NA	1 (0.82%)

Table 5: Tumor stage (p-value = 0.05)

Subtype	Stage	Count
01	Alive	48 (39.3%)
01	Dead	13 (10.7%)
02	Alive	15 (12.3%)
02	Dead	4 (3.28%)
03	Alive	22 (18.0%)
03	Dead	20 (16.4%)

Table 6: Vital status (p-value = 0.01)

3.2.3. LSCC Patients

For patients affected by LSCC (Lung Squamous Cell Carcinoma) it is no longer the size and ability to invade neighboring tissues that is the discriminating factor, as in the case of KRCCC, but the metastatic ability of the tumor and thus the speed of insertion into different organs of the body, not necessarily close to the primary organ. This factor is registered as *Pathologic M* ($p-value = 0.01$) for AJCC staging system and with the *Gender* ($p-value = 0.004$) are key relevant for the patients' stratification. In particular, subtype four identified by the proposed method is more characterized by patients with M0 condition (n:55), that is, with no distant metastasis. This explains the survival curve of subtype four, which has the highest survival probability. In addition, the median age of subtype four turns out to be the lowest observed (*median age = 66.5 years*) against a *median age = 70 years* for subtype one (19), (20).

3.2.4. NSCLC Patients

Within the NSCLC (Non-Small Cell Lung Cancer) cohort, the method identified two distinct subgroups (Fig. 5.f) by integrating genomic, transcriptomic, and clinical data, including curated, myeloid, and immune signatures. The most impactful clinical variables for the classification are *Age*, *Gender*, and *Line of Therapy*, which were tested using the clinical information available for the cohort. Regarding the distribution of conditional *Age* at the identified subtypes, there is a statistically significant difference between the two groups of about ten years. With the first subtype characterized by *median age = 65 years* and the second subtype with *median age = 55 years*.

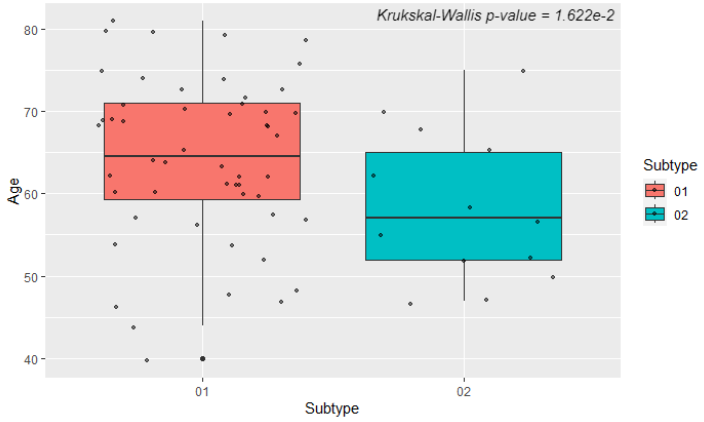


Figure 7: Age distribution for patients' subtypes in NSCLC

The *Gender* variable is an interesting key to interpretation regarding the still-open problem of Sexual Dimorphism (13). As shown in Fig. 8, gender distribution is notable, with subtype two comprising over 85% female patients, At the same time, patients within subtype one are mainly women of an advanced age who have a lower probability of survival than patients in subtype two. The low survival in older women (subtype 1) could be dictated by a menopausal state (13).

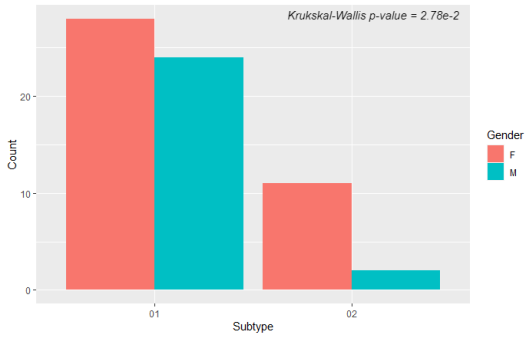


Figure 8: Gender distribution for patients' subtypes in NSCLC

Sex	Type 01 (%)	Type 02 (%)
Female	52.9	85.7
Male	47.1	14.3

Table 7: Gender distribution (percentage) for patients' subtypes in NSCLC

The last statistically significant clinical variable, was the line of therapy, consisting of five different levels depending on the patient's therapy status. As shown in Fig. 9, subtype one includes patients who received all lines of therapy, that is from the first line of therapy to the fifth. Subtype two, on the other hand, which corresponds to the group with a higher probability of survival, consists only of patients who have been treated with the first and second line of therapy. Thus, reviewing the history of clinical variables previously analyzed, it is clear that the subtype with the least likelihood of survival for patients with NSCLC is composed of individuals of women sex, in an advanced age state, and who due to sexual dimorphism resist immunotherapy treatment, which is why subtype one, it is up to the fifth line of therapy.

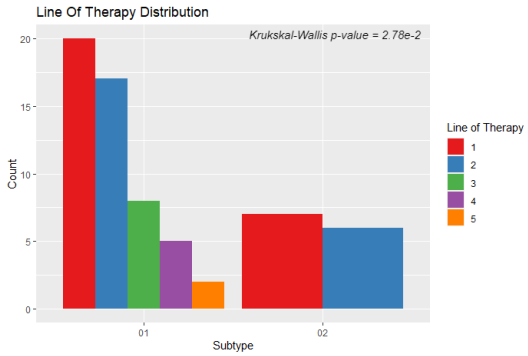


Figure 9: Line of therapy distribution for patients' subtypes in NSCLC

Sex	Type 01 (%)	Type 02 (%)
1	37.3	57.1
2	33.3	42.9
3	15.7	-
4	9.8	-
5	3.9	-

Table 8: Line of therapy distribution (percentage) for patients' subtypes in NSCLC

4. Conclusion

From the proposed work, it is clear the importance of using computational methods based on patient similarity networks (PSNs), and the integration of omics data to achieve statistically and biologically significant stratification of patients within the same cohort affected by complex and heterogeneous diseases. The proposed approach, not only outperforms the SNF algorithm but also brilliantly overcomes data integration with important clinical potential in therapy identification and optimization. From a technical point of view, the proposed methodology out-performed the SNF algorithm obtaining a better stratification of GBM patients in terms of overall survival, identifying three statistically significant subtypes, with the first group of patients oldest and also with the lowest probability of survival compared with the other two groups. Considering NSCLC patients, there is a significant stratification of patients with clinical characterization mainly by sex, and age. Looking at the survival curves it is clear that patients who are within subtype one are the ones who have the lowest probability of survival. Similarly, among the statistically significant clinical variables, patients in subtype one have the highest median age of 65 years. In addition, the variable age suggests that the majority of patients in subtype one are women. From this analysis, the lower probability of survival for older women emerges. Considering the immunotherapy treatment received by the patients on whom data were recorded, and the median age of patients in subtype one, it is likely that this result is related to the phenomenon of sexual dimorphism (13), which refers to biological and physiological differences between the sex of species, including variations in physical characteristics, hormone levels, and immune responses. In particular, in women undergoing menopause, the significant decrease in estrogen can negatively affect the immune response, reducing the effectiveness of immunotherapy, which uses the immune system to fight cancer. The decreasing of estrogen can alter the tumor microenvironment and affect tolerance to side effects of therapy, leading to different responses than premenopausal women and making it more difficult to benefit from treatment (14). The proposed method captures the division by sex of the two subtypes and highlights the biological importance of the sex variable for NSCLC patients, and other relevant factors among the cancer types analyzed. Furthermore, the proposed work also comes as an enrichment to the work done by Wang and colleagues (4) on the cohort of patients with five different types of cancer. In particular, the proposed approach turns out to be able to return not only a statistically significant stratification of patients but also to offer valid clinical enrichment of the analyzed data, an issue that is not addressed by the SNF method in (4). Further enrichment and valuable insight, have also been made concerning the authors' work on the independent Non-Small Cell Lung Cancer (NSCLC) data (9). In that work, the authors studied the stratification ability of a multi-layer network model for response or resistance to immunotherapy, obtaining two statistically significant and representative clusters. However, the weakness of the SNF method, at least on the NSCLC dataset, when considering high-dimensionality data is highlighted in the previous work, i.e. before integrating the different data layers it was necessary to perform a filtering operation on the RNA-count matrix, to extract only the most relevant genes (sorted by degree), using the 10% of genes (hubs) obtained from the construction of a differential co-expression network. With the proposed approach, on the other hand, the *gene redundancy* problem is unexpressed, and thus the method allows all the partial information from the individual networks to be condensed within the multilayer network, proving robust to high dimensionality. The computational approach presented in this work represents a valuable advancement in the field of precision oncology, offering a powerful framework for refining

patient stratification and informing therapeutic strategies. By enabling a more accurate classification of patient subtypes, this methodology enhances clinicians' ability to tailor treatment plans based on the unique genetic and clinical profiles of individuals, thereby optimizing patient outcomes. Such a data-driven, integrative approach is fundamental to the goals of precision medicine, where the understanding of individual molecular and clinical differences can drive more targeted and effective interventions.

References

[1] Petti M, Farina L. Network medicine for patients stratification: From single-layer to multi-omics. WIREs Mech Dis. 2023 Nov-Dec;15(6):e1623. doi: 10.1002/wsbm.1623. Epub 2023 Jun 15..

[2] Erikainen, S., & Chan, S. (2019). Contested futures: envisioning "Personalized," "Stratified," and "Precision" medicine. New Genetics and Society, 38(3), 308–330.

[3] Farina L. Network as a language for precision medicine. Ann Ist Super Sanita. 2021 Oct-Dec;57(4):330-342. doi:10.4415/21.04.08. PMID:35076423.

[4] Wang, B., Mezlini, A., Demir, F. et al. Similarity network fusion for aggregating data types on a genomic scale. Nat Methods 11, 333–337 (2014).

[5] Hofree, M., Shen, J., Carter, H. et al. Network-based stratification of tumor mutations. Nat Methods 10, 1108–1115 (2013).

[6] Bhalla S, et al. Patient similarity network of newly diagnosed multiple myeloma identifies patient subgroups with distinct genetic features and clinical implications. Sci Adv. 2021 Nov 19;7(47):eabg9551. doi: 10.1126/sci-adv.abg9551. Epub 2021 Nov 17. PMID: 34788103;

[7] Cavalli FMG, et al. Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell. 2017 Jun 12;31(6):737-754.e6. doi: 10.1016/j.ccell.2017.05.005. PMID: 28609654; PMCID: PMC6163053.

[8] Pai S, Bader GD. Patient Similarity Networks for Precision Medicine. J Mol Biol. 2018 Sep 14;430(18Pt A):2924-2938. doi: 10.1016/j.jmb.2018.05.037. Epub 2018 Jun 1. PMID: 29860027; PMCID: PMC6097926 PMCID: PMC8598000.

[9] D. Mascolo, L. Farina and M. Petti, "Multi-layer network modelling of genomic and transcriptomic data to investigate the response to checkpoint inhibitors in NSCLC," 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Istanbul, Turkiye, 2023, pp. 2825-2830, doi:10.1109/BIBM58861.2023.10385375.

[10] Ravi, A., et al. Genomic and transcriptomic analysis of checkpoint blockade response in advanced non-small cell lung cancer. Nature Genetics, 55(5), 807–819.

[11] Pj Mucha et al. Community Structure in Time Dependent Multiscale, and Multiplex Networks, Science, 2010 May.

[12] Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. Onnela, J.-P. Community structure in time-dependent, multiscale, and multiplex networks. Science 328, 876-878 (2010).

[13] Klei, S., et al. Sex differences in immune responses. Nat Rev Immunol 16, 626-638 (2016).

[14] Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016 Oct;16(10):626-38. doi: 10.1038/nri.2016.90. Epub 2016 Aug 22. PMID: 27546235.

[15] Swami U, Nussenzveig RH, Haaland B, Agarwal N. Revisiting AJCC TNM staging for renal cell carcinoma: quest for improvement. Ann Transl Med. 2019 Mar;7(Suppl 1):S18. doi: 10.21037/atm.2019.01.50. PMID: 31032299; PMCID: PMC6462602.

[16] Multimodal recurrence scoring system for prediction of clear cell renal cell carcinoma outcome: a discovery and validation study Gui, Cheng-Peng et al. The Lancet Digital Health, Volume 5, Issue 8, e515 - e524

[17] Alinia, S., Ahmadi, S., Mohammadi, Z. et al. Exploring the impact of stage and tumor site on colorectal cancer survival: Bayesian survival modeling. Sci Rep 14, 4270 (2024). https://doi.org/10.1038/s41598-024-54943-8

[18] Thompson RH, Kurta JM, Kaag M, Tickoo SK, Kundu S, Katz D, Nogueira L, Reuter VE, Russo P. Tumor size is associated with malignant potential in renal cell carcinoma cases. J Urol. 2009 May;181(5):2033-6. doi: 10.1016/j.juro.2009.01.027. Epub 2009 Mar 14. PMID: 19286217; PMCID: PMC2734327.

[19] Önal Ö, Koçer M, Eroğlu HN, Yilmaz SD, Eroğlu I, Karadoğan D. Survival analysis and factors affecting survival in patients who presented to the medical oncology unit with non-small cell lung cancer. Turk J Med Sci. 2020 Dec 17;50(8):1838-1850. doi: 10.3906/sag-1912-205. PMID: 32512671; PMCID: PMC7775717.

[20] Pilotto S, Sperduti I, Novello S, Peretti U, Milella M, Facciolo F, Vari S, Leuzzi G, Vavalà T, Marchetti A, Mucilli F, Crinò L, Puma F, Kinspergher S, Santo A, Carbognin L, Brunelli M, Chilosi M, Scarpa A, Tortora G, Bria E. Risk Stratification Model for Resected Squamous-Cell Lung Cancer Patients According to Clinical and Pathological Factors. J Thorac Oncol. 2015 Sep;10(9):1341-1348. doi: 10.1097/JTO.0000000000000628. PMID: 26200453.