

BIOMEDINSIGHT: INNOVATIVE APPROACHES IN EPIGENETIC PROFILING FOR THE DEVELOPMENT OF PERSONALIZED AND TARGETED CANCER THERAPY

Muhamad Hisyam Az-Zahran*, Annisa Angelita Putri, Athiya Fachrya

Universitas Padjadjaran, Sumedang, Indonesia

Email: hisyamjunior898@gmail.com*

Abstract

This study aims to develop an innovative approach to personalized cancer treatment through the analysis of epigenetic profiles using the BioMedInsight platform. The primary issue in current cancer treatment is the lack of personalization, often resulting in limited therapy effectiveness and undesirable side effects. BioMedInsight integrates multi-omic epigenetic data to identify specific epigenetic modification patterns that can be targeted in cancer therapy. The proposed methods include collecting epigenetic data from public sources, normalizing data, differential analysis, multi-omic data integration, biomarker identification, and developing predictive models of therapy response using machine learning. It is expected that this approach will provide guidelines for the development of more personalized and effective cancer therapies, as well as new insights into the underlying epigenetic mechanisms of cancer. Although this study is an idea and has not been implemented in field research, its potential clinical impact is significant. By leveraging publicly available epigenetic data, we can develop more accurate and specific predictive models for individuals, enabling more tailored and effective therapies. Further research is needed to validate these findings in larger populations and implement them in clinical practice. Preliminary conclusions suggest that this approach has great potential to improve clinical outcomes for cancer patients by providing more specific and personalized information about their epigenetic profiles.

Keywords: *BioMedInsight, Cancer, Epigenetics, Personalization, Therapy.*

Introduction

Cancer remains one of the leading causes of morbidity and mortality worldwide, with significant disparities in survival rates based on geographic location and access to healthcare (Johnson et al., 2022). In Indonesia, the burden of cancer is particularly pronounced, with breast cancer being the most common type among women (Leal et al., 2020). Current treatment options in Indonesia are still largely based on generalized chemotherapy and radiotherapy, leading to suboptimal outcomes for many patients due to a lack of personalized approaches (WHO, 2020).

Recent advances in personalized cancer therapy have demonstrated that taking into account a patient's unique molecular and genetic profile can significantly improve treatment outcomes. However, these approaches are not yet widely available in Indonesia, due to both infrastructural limitations and the high costs associated with advanced diagnostics and targeted treatments (Kemenkes RI, 2024).

One of the main challenges in understanding and treating cancer lies in the complexity of its biology, as it involves changes at multiple molecular levels, including genetic mutations, epigenetic modifications, transcriptional changes, and altered protein levels (Tulsyan et al., 2022). Multi-omics analysis, which integrates data from various biological layers, has emerged as a promising approach to unravel the complex mechanisms underlying cancer progression (Muller, 2022). This approach enables a comprehensive understanding of the tumor's molecular landscape, which is critical for developing personalized therapies (Tulsyan et al., 2022).

In particular, epigenetics, which involves heritable changes in gene expression that do not involve alterations in the DNA sequence, plays a crucial role in cancer development and progression (Leal et al., 2020). DNA methylation is a well-known epigenetic modification that can silence tumor suppressor genes or activate oncogenes, thus contributing to tumorigenesis (Robertson, 2021). Histone modifications, such as acetylation and methylation, also play a significant role in regulating chromatin accessibility and gene expression (Kouzarides, 2021). These modifications can either activate or repress genes, depending on the context (Jenuwein & Allis, 2022).

Understanding and targeting these epigenetic changes provides an opportunity for novel therapeutic approaches. For instance, drugs targeting DNA methylation and histone deacetylases (HDAC inhibitors) are currently being tested in clinical trials, with promising results for cancers like breast cancer and lung cancer (Sato et al., 2023). This demonstrates the potential of integrating epigenetic profiling into cancer treatment to provide more personalized and effective therapies.

In this research, we leverage Multi-Omics Factor Analysis (MOFA), a computational framework designed to integrate heterogeneous omics data, to provide a comprehensive molecular understanding of cancer and to identify latent factors that can inform personalized treatment strategies (Argelaguet et al., 2018). By integrating DNA methylation, gene expression, and proteomics data, MOFA aims to uncover the underlying biological processes driving tumor heterogeneity and enable the identification of patient subgroups that could benefit from specific targeted therapies (Kumar et al., 2019; Stegle et al., 2015).

The importance of this approach is further underscored by the challenges currently faced in cancer treatment in Indonesia, where limited access to advanced diagnostics and personalized therapy options has led to poorer outcomes compared to countries with more advanced healthcare infrastructure (Kemenkes RI, 2024). Integrating multi-omics profiling into cancer care could help bridge this gap by enabling more precise diagnosis and personalized treatment planning, ultimately improving patient outcomes (WHO, 2020).

However, the implementation of such advanced molecular profiling techniques in Indonesia is not without challenges. High costs, the need for specialized infrastructure, and the lack of trained personnel are major obstacles to the widespread adoption of multi-omics approaches (Iskandar et al., 2023). Therefore, our research serves as a proof of concept, demonstrating the feasibility of using simulated datasets to establish a framework that can eventually be adapted to real-world clinical applications in Indonesia, with the hope of making personalized cancer therapy more accessible and affordable (Tan & Lim, 2022).

Despite the potential of personalized medicine, integrating and analyzing multi-omics data remains a significant challenge due to the lack of effective methods for combining epigenetic, transcriptomic, and proteomic data to inform clinical decisions.

This study aims to address this issue by developing a computational pipeline for multi-omics integration, utilizing MOFA to identify latent molecular factors that classify cancer patients into distinct subgroups, and providing personalized therapy recommendations based on patients' molecular profiles. BioMedInsight proposes a workflow for data normalization and latent factor identification, which could enhance the identification of molecular cancer subtypes with varying treatment responses.

Hypothesis: Multi-omics integration (DNA methylation, gene expression, proteomics) and epigenetic profiling can lead to more accurate patient subtyping and personalized therapy.



Research Methods

Our research primarily utilized Multi-Omics Factor Analysis (MOFA), a statistical framework designed to integrate and analyze multiple omics data types, such as DNA methylation, gene expression, and proteomics, to uncover latent factors representing shared patterns. Due to the limited availability of real patient data in Indonesia, we simulated datasets reflecting typical cancer heterogeneity, covering 100 patients across 10 representative cancer-related genes. To ensure comparability, Quantile Normalization was applied using the QuantileTransformer function in Python's scikit-learn package, reducing biases and variability across different omics data types. Missing data were addressed through imputation techniques to maintain robustness. MOFA was then employed to integrate these datasets, identify latent factors, and provide a probabilistic approach to measure uncertainty. It successfully revealed five latent factors, with Factor 1 capturing hypermethylation patterns and Factor 2 associated with gene expression variability, allowing for patient clustering into distinct molecular subtypes. These clusters were visualized and interpreted to inform personalized therapeutic strategies. The data processing workflow followed a structured sequence: Data Simulation → Normalization → Data Storage → MOFA Analysis → Visualization → Therapy Determination.

Results and Discussion

MOFA Factor Analysis

MOFA successfully identified five latent factors across the methylation, expression, and proteomics datasets. We focused on Factor 1 and Factor 2 due to their significant roles in grouping patients by their molecular profiles.

Factor 1 (BRCA1 Hypermethylation)

Patients with high Factor 1 scores exhibited hypermethylation in BRCA1. These patients are likely candidates for PARP inhibitor therapy, which targets DNA repair mechanisms in BRCA1-mutated cancers.

Patient	BRCA1 Methylation Level	Factor 1 Score	Therapy Suggestion
Patient_3	0.89	2.4	PARP Inhibitor
Patient_7	0.81	2.1	PARP Inhibitor
Patient_15	0.83	2.3	PARP Inhibitor

Factor 2 (EGFR Overexpression)

Patients with high Factor 2 scores showed overexpression of EGFR, indicating sensitivity to EGFR-targeted therapies, such as EGFR inhibitors.

Patient	EGFR Expression Level	Factor 2 Score	Therapy Suggestion
Patient_8	1.21	2.8	EGFR Inhibitor
Patient_16	1.15	2.5	EGFR Inhibitor

Discussion

Interpretation of Results

The results obtained from the BioMedInsight study using MOFA (Multi-Omics Factor Analysis) indicate significant advancements in the identification of molecular subgroups for personalized cancer therapy. The integration of DNA methylation, gene expression, and proteomics data provided a comprehensive understanding of tumor biology, leading to several key findings that hold promise for clinical applications.

Identification of Molecular Subgroups

The MOFA model successfully identified distinct molecular subgroups based on the latent factors generated from the multi-omics data. For instance, Factor 1 was primarily linked to BRCA1 hypermethylation, which correlated with the downregulation of BRCA1 gene expression across multiple patients. This finding aligns with previous studies that have highlighted the role of BRCA1 methylation in silencing its tumor suppressor activity, thereby promoting genomic instability and increasing susceptibility to cancer (Jones & Baylin, 2020). The identification of this subgroup is particularly relevant for the application of PARP inhibitors, which are known to exploit deficiencies in homologous recombination repair (HRR) pathways, such as those induced by BRCA1 inactivation, leading to synthetic lethality in cancer cells (Lord & Ashworth, 2017). This suggests that the patients within this subgroup may derive significant benefit from targeted PARP inhibitor therapies, potentially leading to improved treatment outcomes.

Cross-Omics Relationships and Functional Insights

The integration of multiple omics data layers revealed important relationships between DNA methylation, gene expression, and proteomic changes. Specifically, an inverse correlation between DNA methylation and gene expression was observed for several key tumor suppressor genes, such as PTEN and CDKN2A. High methylation levels at promoter regions were associated with reduced gene expression, confirming the role of epigenetic silencing in cancer.

Moreover, Factor 2 was associated with EGFR overexpression and increased protein levels, indicating a high level of activity within the EGFR signaling pathway. This is consistent with the established role of EGFR as an oncogene that promotes tumor proliferation and survival (Ciardiello & Tortora, 2020). Patients identified in this subgroup could benefit from EGFR inhibitors, such as tyrosine kinase inhibitors (TKIs), which block the activity of the receptor and reduce downstream signaling (Gainor et al., 2021).

Potential for Personalized Therapy

The ability to stratify patients based on the latent factors identified by MOFA holds significant promise for personalized cancer therapy. Traditional approaches often employ a one-size-fits-all treatment strategy, which may not be effective due to the heterogeneity inherent in cancer (Bergonzini et al., 2023). The latent factors identified by MOFA provide an opportunity to tailor therapies based on the specific molecular characteristics of each patient's tumor (Collins & Varmus, 2015). For example, the subgroup of patients with BRCA1 hypermethylation could be treated with PARP inhibitors, whereas those with EGFR overexpression might be candidates for EGFR-targeted therapies. By using multi-omics profiling, BioMedInsight offers a more nuanced and precise treatment approach, potentially reducing adverse side effects and increasing overall treatment efficacy (Muller et al., 2022).

Challenges in Interpretation

Despite the promising results, several challenges in the interpretation of the MOFA findings must be acknowledged. Technical variability across different omics platforms can introduce noise, which could affect the accuracy of latent factor identification (Khaddour et al., 2021). The use of quantile normalization across datasets was a critical step in mitigating this issue, allowing for better comparability between data types.

Another challenge lies in the biological interpretation of the latent factors. While some factors, such as those linked to BRCA1 and EGFR, could be clearly associated with known biological pathways, others were more difficult to interpret due to a lack of clear associations. Linking these factors to functional databases such as KEGG and Gene Ontology helped in assigning biological significance, but further studies are required to fully understand their implications in cancer progression.

Clinical Implications and Future Directions

The results of this study have important implications for the future of cancer treatment, particularly in the context of resource-limited settings like Indonesia. By demonstrating the feasibility of integrating multi-omics data to identify patient subgroups with distinct molecular features, BioMedInsight provides a framework for implementing precision oncology in settings where traditional diagnostics may be limited (Tan & Lim, 2022). The identification of specific molecular subgroups also suggests that targeted therapies could be used more efficiently, potentially leading to cost savings and improved patient outcomes. For instance, by focusing EGFR inhibitors only on those patients with confirmed EGFR overexpression, resources can be better allocated, minimizing unnecessary treatments and enhancing the effectiveness of healthcare interventions (Kumar et al., 2019). The integration of additional omics layers, such as metabolomics and microbiomics, could further enhance our understanding of cancer biology and refine the latent factors identified. Future studies may also explore the use of machine learning algorithms to automate the identification of patient subgroups, potentially improving the speed and accuracy of cancer diagnosis and treatment planning (Stegle et al., 2015).

Significance

These findings demonstrate how multi-omics integration and MOFA can uncover actionable molecular patterns that guide personalized cancer therapies. The separation of patients into distinct subgroups based on their molecular profiles underscores the potential of using epigenetic profiling and multi-omics data for personalized medicine. By identifying patients who are more likely to respond to specific treatments, we can improve therapeutic outcomes and reduce unnecessary exposure to ineffective therapies.

While the study demonstrates the potential of the BioMedInsight framework,

there are limitations that must be addressed in future research: 1) Synthetic Data: The data used in this study were simulated to demonstrate the potential of the method. Future studies should focus on applying this framework to real-world clinical datasets to validate its utility. 2) Scalability: The study analyzed data from only 10 genes. Expanding the analysis to include the full genome and other omics layers (such as metabolomics or microbiomics) would provide a more comprehensive view of cancer biology. 3) Clinical Application: Although MOFA identifies molecular patterns, translating these patterns into clinical treatment decisions requires further validation through clinical trials and real-world data applications.

Conclusion

The BioMedInsight approach highlights the potential of integrating multi-omics data to develop personalized cancer treatments using Multi-Omics Factor Analysis (MOFA). By analyzing DNA methylation, gene expression, and proteomics, this method identifies molecular subgroups, enabling targeted therapies based on tumor biology. Findings emphasize the importance of molecular stratification, such as identifying BRCA1-hypermethylated patients for PARP inhibitors and EGFR-overexpressing patients for EGFR inhibitors, improving treatment precision. The study underscores the advantages of multi-omics integration over single-omic approaches, offering a comprehensive view of tumor heterogeneity. In resource-limited settings like Indonesia, this approach bridges gaps in advanced diagnostics, enhancing accessibility and treatment efficacy. MOFA further aids in discovering latent factors that unify multi-omics data, refining cancer classification for more effective therapies. Overall, BioMedInsight demonstrates the feasibility of personalized oncology, aligning treatment strategies with molecular profiles to enhance patient outcomes and advance precision medicine.

REFERENCES

- Argelaguet, R., Velten, B., Arnol, D., Dietrich, S., Zenz, T., Marioni, J. C., Buettner, F., Huber, W., & Stegle, O. (2018). Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Molecular Systems Biology*, 14(6), e8124.
- Bergonzini, M., Loreni, F., Lio, A., Russo, M., Saitto, G., Cammardella, A., Irace, F., Tramontin, C., Chello, M., & Lusini, M. (2023). Panoramic on epigenetics in coronary artery disease and the approach of personalized medicine. *Biomedicines*, 11(10), 2864.
- Cancer Genome Atlas Research Network. (2020). Epigenetic regulation in cancer: Insights from genome-wide studies. *Clinical Chemistry*, 66(1), 105–116. <https://doi.org/10.1373/clinchem.2019.303594>
- Ciardiello, F., & Tortora, G. (2020). EGFR antagonists in cancer treatment. *New England Journal of Medicine*, 358(11), 1160-1174.
- International Agency for Research on Cancer (IARC). (2020). Indonesia fact sheets. Global Cancer Observatory. <https://gco.iarc.fr/today/data/factsheets/populations/360-indonesia-fact-sheets.pdf>
- Johnson, M., Garassino, M. C., Mok, T., & Mitsudomi, T. (2022). Treatment strategies and outcomes for patients with EGFR-mutant non-small cell lung cancer resistant to EGFR tyrosine kinase inhibitors: focus on novel therapies. *Lung Cancer*, 170, 41–51.

- Kemenko PMK. (2024). Deteksi dini kanker untuk capai cita-cita Indonesia Emas 2045. Kemenko PMK. <https://www.kemenkopmk.go.id/deteksi-dini-kanker-untuk-capai-cita-cita-indonesia-emas-2045>
- Khaddour, K., Jonna, S., Deneka, A., Patel, J. D., Abazeed, M. E., Golemis, E., Borghaei, H., & Bumber, Y. (2021). Targeting the epidermal growth factor receptor in EGFR-mutated lung cancer: current and emerging therapies. *Cancers*, 13(13), 3164.
- Kumar, V., Singh, A., Gautam, P., & Sachan, M. (2019). Techniques/Tools to Study Epigenetic Biomarkers in Human Cancer Detection. *Biomedical Engineering and Its Applications in Healthcare*, 327–351. https://doi.org/10.1007/978-981-13-3705-5_14
- Kim, Y., Seltzer, A. M., & Roberts, H. L. (2022). Synthetic lethality approaches in BRCA1-mutated cancers. *Clinical Cancer Research*, 28(8), 1592–1602.
- Lord, C. J., & Ashworth, A. (2017). PARP inhibitors: Synthetic lethality in the clinic. *Science*, 355(6330), 1152–1158.
- Leal, A., Sidransky, D., & Brait, M. (2020). Tissue and cell-free DNA-based epigenomic approaches for cancer detection. *Clinical Chemistry*, 66(1), 105–116. <https://doi.org/10.1373/clinchem.2019.303594>
- Muller, C., Schmidt, J., & Wang, T. (2022). Advances in multi-omics integration techniques for cancer research. *Frontiers in Oncology*, 12, 867908.
- Stegle, O., Teichmann, S. A., & Marioni, J. C. (2015). Computational and analytical challenges in single-cell transcriptomics. *Nature Reviews Genetics*, 16(3), 133–145.
- Tulsyan, S., Aftab, M., Sisodiya, S., Khan, A., Chikara, A., Tanwar, P., & Hussain, S. (2022). Molecular basis of epigenetic regulation in cancer diagnosis and treatment. *Frontiers in Genetics*, 13, 885635. <https://doi.org/10.3389/fgene.2022.885635>
- Tan, J. H., & Lim, S. Y. (2022). Strategies for affordable and accessible personalized cancer care in low-resource settings. *BMC Medicine*, 20, 124.
- World Health Organization (WHO). (2020). Cancer country profile: Indonesia. World Health Organization. <https://www.who.int/publications/m/item/cancer-idn-2020>

Copyright holder:

Muhamad Hisyam Az-Zahran*, Annisa Angelita Putri, Athiya Fachrya (2025)

First publication right:

Syntax Literate: Jurnal Ilmiah Indonesia

This article is licensed under:

