Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 27 June 2026
Revised 27 July 2026
Accepted 10 Aug 2026

Natural Resources for Human Health 2026, Vol. 6 (18s): 426–438

KEYWORDS:
Artificial intelligence, cancer stem cells, deep learning, drug discovery, explainable AI, multi-omics integration, PLS-SEM, and precision oncology.

AI-Driven Multi-Omics Integration for Targeted Drug Discovery Against Cancer Stem Cells: A Genomic Perspective

Anwita Dhawan
San Francisco State University, San Francisco, Dhawananwita9@gmail.com

ABSTRACT: Cancer stem cells (CSCs) are responsible for the initiation of tumors, their relapses and drug resistance to a very high extent, and the molecular heterogeneity between these cells is rarely identified by single-omics analysis. Artificial intelligence (AI) can be very useful for the integration of genomic, transcriptomic, epigenomic, and proteomic data, but it is less known if scientists will value this research. This study examines how much researchers perceive the effectiveness of AI-driven multi-omics integration for the clinical applicability in CSC-directed drug discovery. A structured questionnaire of 15 items was administered to 63 respondents from [bioinformatics, biotechnology, oncology and allied fields], and the data were analyzed using partial least squares structural equation modelling (PLS-SEM) in SmartPLS 4. All constructs showed acceptable internal consistency (Cronbach's α = 0.785–0.813; composite reliability = 0.853–0.869; AVE = 0.542–0.576) and no collinearity concerns (VIF ≤ 3.307). AI-driven multi-omics integration was strongly associated with perceived classification capability (β = 0.833, R^2 = 0.694) and with explainability and clinical applicability (β = 0.875, R^2 = 0.765). However, since heterotrait–monotrait ratios are above 1.0, it reveals that respondents did not clearly separate the three constructs, suggesting they perceive AI integration, classification and explainability as a single, unified value proposition. The findings indicate that researchers have strong confidence in integrated AI approaches for precision oncology, while underlining the need for sharper construct measurement and empirical benchmarking of AI models on real multi-omics data.

I. INTRODUCTION

Cancer is a global burden of disease causing a worldwide health impact and the researchers in the medical field have been actively working on developing better therapeutics and improving the current treatments. The United States projected 2,001,140 new cancer cases and 611,720 cancer deaths for 2024. (Velmurugan et al., 2025). Though a lot of breakthroughs in the field of oncology have helped lower the mortality rates, a permanent intervention for metastasis and drug resistance is yet to be discovered. Researchers have detected that the cells that actually cause the tumor to metastasize and develop drug resistance and are known to be cancer stem cells. Cancer stem cells are seen as a major obstacle in research because they allow the tumor to outlive harsh treatments. The relapse of tumors and the development of new ones majorly occur due to these cells. Hence analyzing and researching these cells can help scientists understand how these cells attack the immune system and the interactions occurring in tumor microenvironment. These studies are useful for scientists to understand the drawbacks of current conventional treatments. (El-Tanani et al., 2025). Unlike the bulk of differentiated tumor cells, CSCs possess self-renewal capacity, multipotency, and quiescent metabolic states that render them intrinsically resistant to conventional chemotherapy and radiotherapy. Hence, targeting cancer stem cells and destroying them can prevent situations like tumor recurrence, metastasis, and body resistance towards treatments. This can be done by investigating signaling pathways for cancer stem cells and analyzing the environment that supports tumor growth. Eventually, a multidisciplinary approach, using tools like single cell sequencing and incorporating artificial intelligence can help develop therapeutic approaches that are beyond traditional methods like surgery, chemotherapy and radiotherapy. (Lee et al., 2025) Cancer stem cells hold the capabilities of bypassing the human body's immune system due to its symbiotic relationship with tumor microenvironment as shown in figure

1. This mutual interaction causes cancer progression and therapeutic resistance. Researchers are actively focusing on building a targeted treatment and eliminating the cellular interaction (symbiosis) between cancer stem cells and tumor microenvironment remains the key idea. (Mathan & Singh, 2025).

An approach which can be beneficial for achieving the goals of precision medicine can be integrating artificial intelligence, multi omics, and computational frameworks together. Artificial Intelligence can have a ground-breaking impact in Cancer research by analyzing the biological data and studying various aspects of tumor behavior and cellular functions. Combining the research with multi omics can help researchers comprehend this better. Moreover, computational frameworks can help identify some gaps within cancer stem cells by identifying regulatory networks through methods like deep learning and data fusion to study huge genomic and proteomic datasets. Such an integrated approach can be the future of more successful therapies in cancer. (Li et al., 2023) A scientific approach of multi-omics including genome, transcriptome and proteome could be a revolutionary approach in the field of drug discovery. Technological advancements in these fields enable the development of spatial atlases. Such an approach allows us to better understand the cellular interactions occurring due to human diseases. Such modern techniques can help identify different molecular profiles within a single tumor mass, hence improving and building more targeted cancer therapies. This way Artificial Intelligence and integrated data science can transform modern medicine. (Kant et al., 2025)

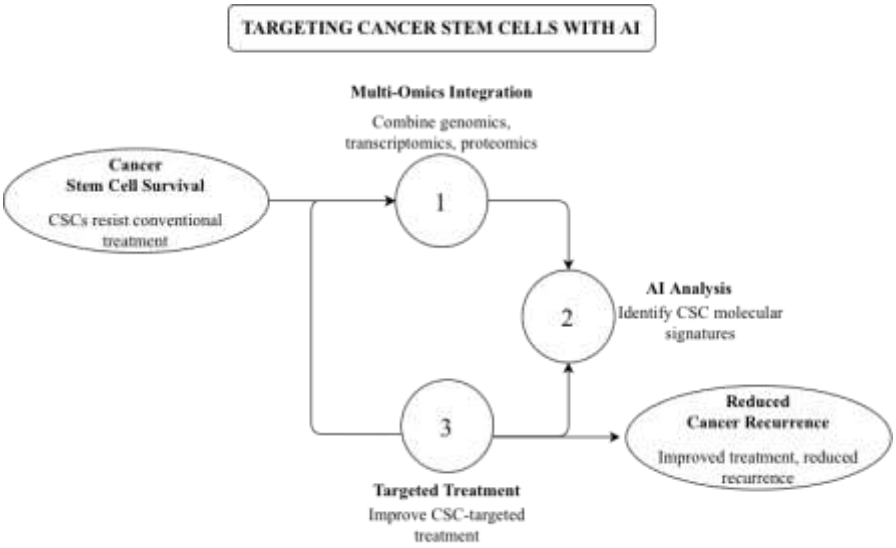


Figure 1 : Targeting Cancer Stem Cells With AI

Artificial intelligence can help predict cancer in early stages, and they can also help us to predict which patients will respond to which kind of therapy. This can be done by using biological insights, clinical records, and medical imaging. AI can eventually create a digital twin for that specific for each patient. Such a prototype can help understand the prognosis of the disease and eventually help doctors understand the treatment strategies that would work for the patient. A way to better understand complexity in biology would be to analyze graph neural networks and transformers. Such advanced machine learning architectures would let researchers study the untraditional relationship between genomic and proteomic layers. This is something traditional statistics usually miss. Such an approach where artificial intelligence is combined with multi-omics data can transform drug discovery and improve precision oncology. (Liu et al., 2024) From the discoveries done in the laboratories to how they are applied in a clinical setting would require an integration that could expedite a shift towards modern medicine. Such integration would include deep learning. Deep learning can help build better therapeutic targets by studying vast biological data sets accurately. Advancements in drug discovery for serious diseases like cancer can be made through merging these technologies. (F. Liu et al., 2026) Emergence of new targeted therapies has often had many limitations due to CSCs-mediated therapeutic resistance. Discoveries have been done to understand pathways to target cancer stem cells, some examples of these pathways include WNT/β-Catenin, hedgehog, Notch, NF-ϰB, JAK/STAT, TGF-β, PI3K/AKT, PPAR. After all the advanced technologies and discoveries, a treatment method to destroy Cancer stem cells still remains undetermined. Currently targeting the biomarkers of Cancer Stem Cells and understanding its regulatory mechanisms remain the idea for developing successful

treatments. (Y. Liu et al., 2026b) Chromatin Remodeling is one another reason that makes cancer stem cells resistant to drugs. Precision medicine for malignant tumors can be achieved through targeted epigenetic therapies. Analyzing biomarkers is a way to predict how patients would react to these therapies. Blending modern technologies like AI Driven drug repositioning could be a way to build more targeted treatments for malignancies. Integration of such modern technologies would help achieve better clinical outcomes due to more accurate patient stratification. (Salvati et al., 2025) A way cancer stem cells are able to survive traditional chemotherapies is genetic mutations. Since cancer stem cells are genomically unstable they undergo a lot of mutations and hence clonal evolution occurs which facilitates epithelial to mesenchymal transition. This allows the cells to overcome traditional treatments. Furthermore, biological markers and signaling pathways like PARP inhibitors expedite the ability of a tumor to surpass medical treatments. With multi-omics integration, these cell types can be targeted and eliminated. (Alizadeh et al., 2025). Cancer can be detected in early stages and that significantly improves the recovery rate by traditional therapies like surgeries and chemotherapy. But the relapse of cancer and recurrence of tumors is very common. Such relapse usually happens due to complex genomic mutations and A tumor microenvironment that prevents the medicine to act positively on the body. Such challenges can be prevailed through an integration of artificial intelligence as well as multi-omics data, because they can help in specifying the particular molecular targets that are causing the cancer stem cells to persist even after strong therapeutics. Improvement in precision medicine and targeted drug discovery can be achieved through these computational approaches. (Chu et al., 2024) One of the reasons for recurrence of cancer is mechanisms like DNA repair and metabolic flexibility in cancer cells. These mechanisms can be better targeted through immunotherapies when the tumor microenvironment is analyzed and integrated with multi-omics data. (Ghosh et al., 2025)

Multi omics integration would improve the accuracy prognosis for a patient, hence moving towards advanced precision oncology. A standard treatment might not work for all patients since all human bodies react differently. Here, computational work can be beneficial because it can read massive biological datasets and then develop personalized treatments for each patient. Such way, scientists would move one step forward towards targeted drug discovery. (Biswas et al., 2026). With emerging modern oncology research, Cancer can be characterized as a disturbance in multiple biomolecular layers like the genetics and proteins. This study was done to a vast amount of data; artificial intelligence can browse. Rather than studying graph neural networks, transformers and vast datasets through statistical methods, AI models can do this research and generate more accurate relationships within the multi omics datasets more rapidly as shown in figure 2. Cancer has various subtypes, and the computational tools can recognize different therapeutic targets according to the focus of infection. These tools take up material from the various data repositories and can predict prognosis of treatment, patient outcomes, and reactions to different therapies. These computational tools also expedite the discovery of treatments that specifically target cancer stem cells. The future of personalized medicine lies in the convergence of machine learning and genomic perspective. (N. Biswas & Chakrabarti, 2020)

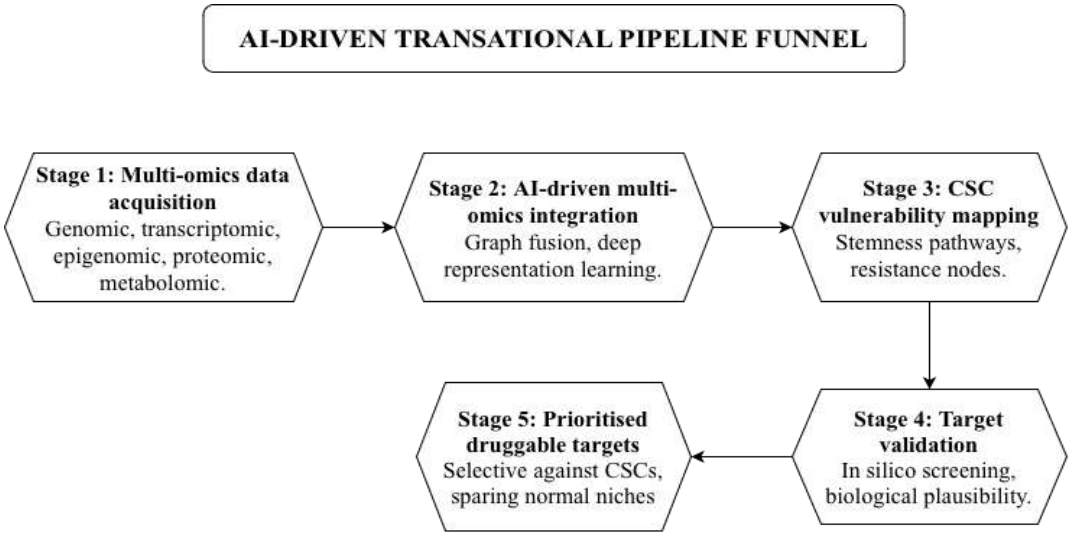


Figure 2: AI Driven Translational Pipeline Funnel

A standardized methodology where researchers can use massive datasets to build personalized therapies would be a revolutionary breakthrough in the field of medicine. The idea of using AI to understand the complex biological data for cancer treatment is something researchers all over the world are working on, but personalized medicine is something that yet looks like a speculative idea. This is because there is a lack of integration between the various cellular layers like genetics and protein behavior. To overcome these drawbacks, a genomic-centered architecture is being built so that it can act as a bridge between computational predictions and practical medical applications.

II. LITERATURE REVIEW

Cancer deaths and new cases have been predicted to rise by 47% worldwide by the year 2040, with some countries facing an increase up to 95% (Albaradei, 2025b). Therefore, it is extremely important to build targeted therapeutics and precision medicine in cancer. Some technologies like the modern Single-cell multi-omics have an important role in the improvement of modern cancer research, and its evolution directly impacts translational research and targeted drug discovery. Cancer stem cells were previously understood as a permanent cell type that cannot change its traits and exists in tumor tissues but with progress in throughput, resolution and data integration, cancer stem cells are now understood as cells that can change their identity and are in a flexible functional state i.e. Differentiated cells can revert into cancer stem cells and cancer stem cells can mature to be differentiated tumor cells. The use of these technologies also helps in mapping cell lineages and building atlases that are tissue specific. Bioinformatic tools and mathematical modelling also hold a lot of importance in understanding the vast amount of biological data. Such an integrated multi-omic approach can further interpret the intricacies of malignant cells. (Baysoy et al., 2023) Cancer stem cells are strongly associated with tumor heterogeneity, drug resistance and disease relapse, and these are the strongest challenges for treating cancer. Though techniques like genomic sequencing make understanding cancer stem cells better at molecular level, measuring those biological traits accurately is still a challenging task with existing computational tools. One of these tools is the one-class logistic regression model. It is used to measure stemness and to make predictions for the patients but recently this tool has been inaccurate and such issues with reliability suggests the need for more AI-driven multi-omics integrated models that can properly identify therapeutic targets. (Z. Liu et al., 2022) These models can be made by combining different biological datasets like genomics and proteomics. Such combinations create a standard framework with which scientists are able to identify the precise molecular structure of tumor cells. The goal of building personalized medicine that is customized for each patient's genetic profile can be reached when the behavior of cancer stem cell is well understood and there are drug targets that have a high rate of therapeutic efficiency as shown in figure 3.(Catalano et al., 2025) Integrating genomics, transcriptomics and the tumor microenvironment is also a way for scientists to identify different signaling pathways so that they can understand how do cancer stem cells actually resist treatments. To enhance patient stratification and build more advanced targeted therapies, future research should focus on using basic science mechanisms to understand clinical applications. This gap in biology can be filled by incorporating high resolution spatial tools. Another approach towards building novel biomarkers is utilizing bioinformatic platforms and public database repositories. (Shrestha et al., 2025) With the help of modern technology and tools scientists are now able to identify cancer stem cells as dynamic identities i.e. stemness is no longer considered as a property of stem cells rather it is seen as a fluid functional state. The way these cells adapt to different environments is understood by 2 different methods, trajectory inference and spatial transcriptomics. Trajectory inference is a computational method used primarily in bioinformatics, and spatial transcriptomics is a way to make a blueprint of gene expression. Weaknesses in traditional treatments are often better understood by such a combination of AI and multi-omics data. With the help of these technologies, scientists aim to create treatment methods that can overcome the obstacles current traditional therapies face, such as cellular plasticity and cancer recurrence. (Huang et al., 2025).

Previously, many discoveries using multi-omic approaches have been made, and they proved to be helpful in oncological studies. One such discovery was when scientists used multi-omic analysis to identify genes that regulate stem cell behavior. A set of sixty-five genes was found to be linked with stemness of tumor cells and when the study was carried out further, it was identified that high properties of stemness indicate poorer prognosis. The LDLR gene was found to be one of the primary genes that connect the trait of stemness in a cell according to the tumor microenvironment. Multiple laboratory experiments like single cell sequencing were done to confirm this finding. This example further elucidates how utilizing AI in medicine and scientific research can help to build targeted therapies. (Yuan et al., 2023)A consistent thread across this body of work is the recognition that no single omics layer adequately explains CSC behavior; rather, the field is moving toward integrative frameworks capable

of reconciling sparse, noisy, and differently scaled data types into a coherent molecular portrait. This has created fertile ground for computational integration strategies, setting the stage for the AI-driven approaches examined in the following theme.

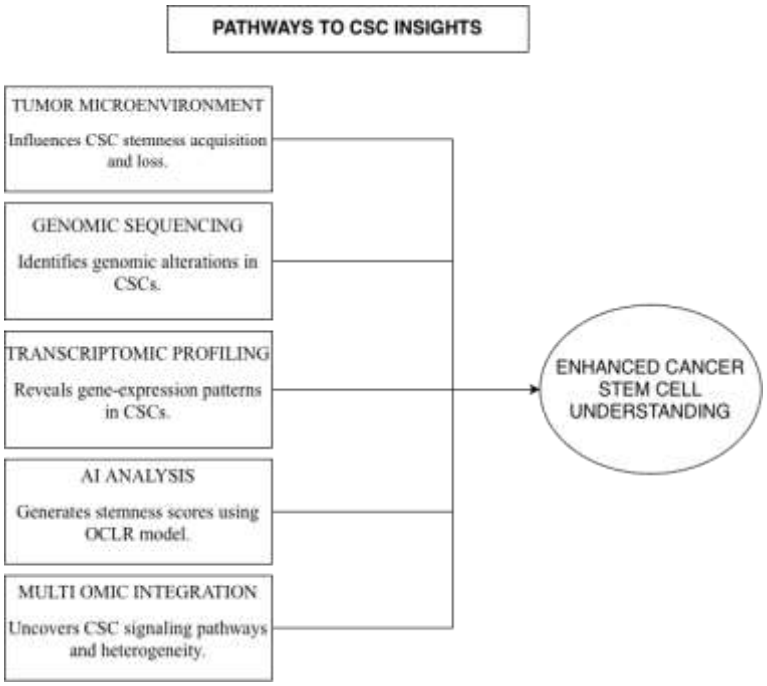


Figure 3: Pathways to CSC Insights

Artificial Intelligence has a prominent role in the pharmaceutical industry as well. During a global health crisis, when fast and efficient therapeutic development is especially required, AI can have a huge role in drug repurposing and building better clinical trial designs. AlphaFold and graph neural network are some of the computational tools that can facilitate patient enrolment and study them in real time to build better clinical study protocols. Prediction of toxicity in clinical trials can be better understood if AI is incorporated with emerging fields like quantum computing and biosensors, though the quality of data and ethical responsibility still remains a hurdle in this field of research. (Vora et al., 2025) Achieving precision drug discovery and being accurate in cancer diagnosis at early stages is an essentially important step towards modern oncology. Utilizing tools like generative AI and deep learning would enable alignment of different biological layers like genomics, proteomics, and metabolomics. From computational tools to scientific research and finally their clinical application, there are gaps that need to be addressed. Once scientists are able to form a framework that finally bridges these gaps, development of individualized cancer treatments would be possible through high-tech data science. (Sikder et al., 2023). Multi-omics dataset is vast and very diverse, hence studying multiple molecular layers allows scientists to understand the complicated relationship between various biological systems. But reviewing enormous amounts of data is difficult, therefore integrating advanced machine learning algorithms would be valuable in interpreting results from those huge datasets. Such computational tools could assist healthcare professionals in making a more accurate disease diagnosis and developing highly targeted therapeutic strategies. (Mukherjee et al., 2024) AI and multi-omics integration can produce some of the most groundbreaking results in oncological research. Multi-omics approach allows to connect information from various genes, proteins and metabolites. This further allows researchers to recognize biomarkers that cause the formation and progression of tumors. Major genes causing tumor progression like TP53 and ERBB2 can be identified, and the tumor microenvironment around these genes can be better understood via these multi-omic approaches. Since multi-omic approach involve complex datasets, managing them using appropriate tools is important and here, AI algorithms are essential because the help in interpreting the vast datasets and ultimately understanding the microenvironment around tumor cells. (Albaradei, 2025). Modern technologies have groundbreaking results in improving the success rate of drug development. Advanced imaging and nanoscale characterization have allowed scientists to better understand cellular interaction and drug disposition, whereas high-throughput "omics" technologies allow scientists to recognize biomarkers and therapeutic targets. For complete ultization of multi-omics data, use of AI particularly machine learning and deep learning is especially important because they enable the analysis of huge datasets. Integration of these

technologies and tools helps researchers identify the most feasible drug targets in a relatively short span of time. Ultimately, this integration would expedite the process of creating highly targeted and precise medical treatments. (Chakraborty et al., 2026). By utilizing the multiple biological layers and synthesizing multi-omics data, scientists are able to predict patient responses, and this helps healthcare professionals in selecting the combination of therapies that would work effectively for a patient. Here explainable AI can have a very vital role in ensuring that the complex amount of data and outputs are converted into an implementable biological way. Though there are a lot of advantages in this field, data heterogeneity and privacy concerns still remain a hurdle. Ultimately, the aim is to bridge the gap between computational models and real-world clinical applications. (Wang et al., 2026)

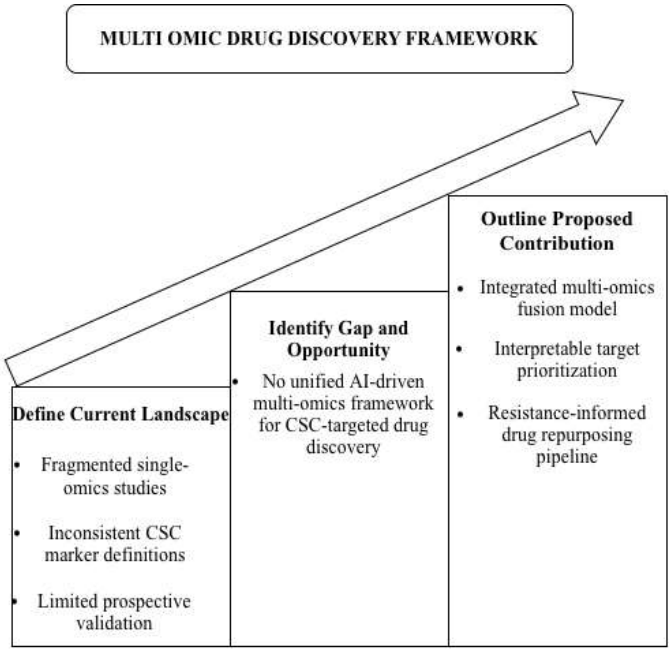


Figure 4: Multi-Omic Drug Discovery Framework

A growing subset of studies uses AI-integrated multi-omics pipelines to identify resistance biomarkers prospectively, enabling stratification of patient subgroups likely to respond to CSC-targeted agents. Work on drug repurposing and combination therapy design increasingly draws on these integrated signatures to nominate synergistic regimens capable of co-targeting bulk tumor cells and their stem-like reservoirs. Despite this progress, the literature also flags persistent translational gaps: limited prospective clinical validation, inconsistent CSC marker definitions across studies, and a scarcity of standardized benchmarking datasets for AI models in this domain. These gaps collectively frame the novelty space that the present study seeks to address as shown in figure 4.

III. RESEARCH METHODOLOGY

A. Research Design This study adopted a quantitative, cross-sectional survey design to examine how researchers and practitioners perceive AI-driven multi-omics integration for CSC-targeted drug discovery. Partial least squares structural equation modelling (PLS-SEM) was chosen over covariance-based SEM because the study is exploratory and prediction-oriented, the sample is modest, and Likert-scale data rarely meets multivariate normality assumptions (Hair et al., 2018). Analysis was carried out in SmartPLS 4 as shown in Figure 5

Natural Resources for Human Health
eISSN: 2583-1194

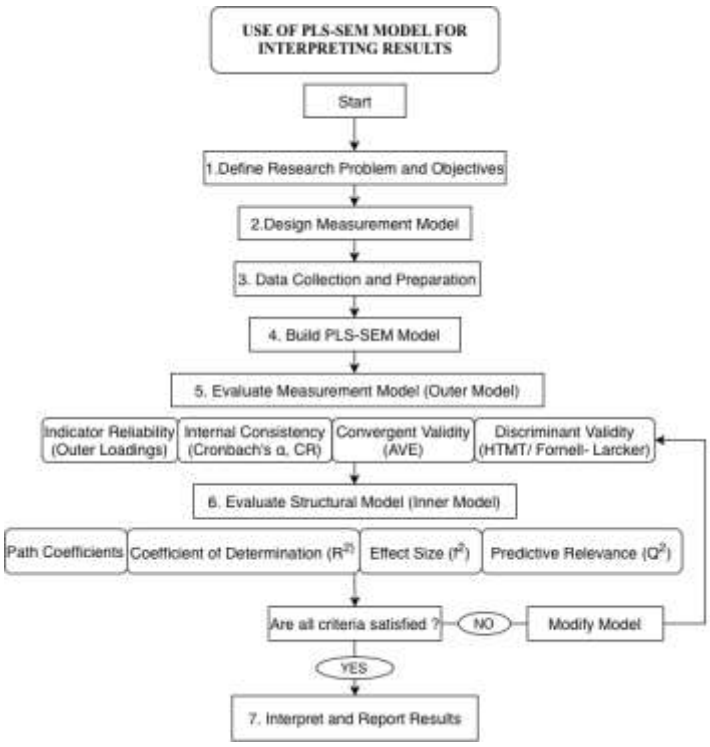


Figure 5: Use of PLS-SEM Model for Interpreting Results

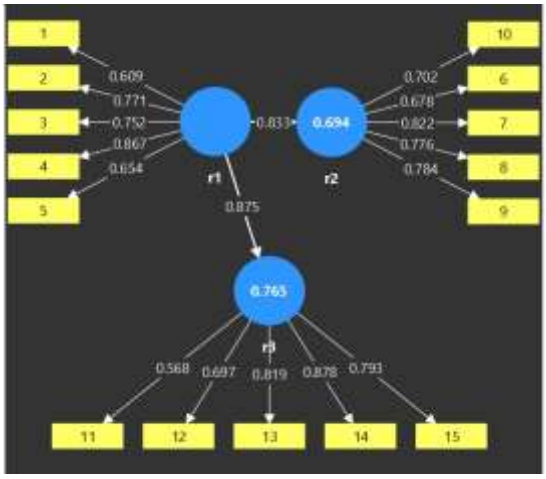


Figure 6: PLS-SEM measurement and structural model

The PLS-SEM measurement and structural model is shown in figure 6. It represents a standard output digaram from the software, SmartPLS. The diagram contains three structural models (the blue circles) and three latent constructs (r1, r2, r3) wherein r1 is an independent latent construct and predicts both of others. r1 to r2 has a path coefficient of 0.833 and r1 to r3 has a path coefficient of 0.875. Both have very strong effects and each of the latent constructs is measured by five survey items. The yellow boxes in the diagram represent the measurement model. The arrows point from the construct to its items hence it is understood that each item is a reflection of the construct. The numbers on these arrows show how strongly each item is represented. These numbers are outer loadings and the usual benchmark is 0.70 or higher.

B. Conceptual Model and Hypotheses

The model bridges three latent variables derived from the research questions. The first is AI-Driven Multi-Omics Integration (AMI, R1), its reflets how the researchers perceive the effectiveness of combining genomic, proteomic, transcriptomic and

epigenomic data through AI to identify the specific targets for CSCs. The second is ML/DL Classification Capability (MLC, R2), it reflects the perceived ability of machine learning and deep learning architectures to identify CSCs among a whole lot of tumor cells. Finally, the third is Explainability and Clinical Applicability (XCA, R3). It reflects how the value of explainable AI (XAI) in making predicted targets interpretable and clinically usable is perceived by researchers. Integration is treated as the foundation: richer, integrated data should strengthen confidence both in classification and in interpretable, clinically usable outputs. The following hypotheses were tested:

- H1: AI-driven multi-omics integration has a significant positive effect on perceived ML/DL classification capability for distinguishing CSCs from bulk tumour cells.
- H2: AI-driven multi-omics integration has a significant positive effect on the perceived explainability and clinical applicability of predicted CSC therapeutic targets.

C. Instrument Development

A structured questionnaire of 15 items (five per construct) was developed from the literature on multi-omics integration, AI in oncology, and explainable AI [cite sources]. Items were measured on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree) [confirm scale]. The draft was reviewed by [number] subject experts for content validity and pilot-tested with [number] respondents, after which wording was refined for clarity. The questionnaire also recorded respondent profiles: academic level, discipline and experience with omics or AI tools.

D. Sampling and Data Collection

Purposive sampling was used to reach respondents with relevant knowledge, including [undergraduate and postgraduate students, researchers, bioinformaticians and clinicians in oncology-related fields]. Data were collected through an online form between 09/17/2026 and 09/23/2026. After screening for incomplete and straight-lined responses, 63 valid responses were retained. This surpasses the minimum of 55 recommended by a G*Power analysis (effect size $f^2 = 0.15$, $\alpha = 0.05$, power = 0.80, one predictor) and the ten-times rule for the largest number of paths into a construct. Participation from respondents was voluntary and completely anonymous. Moreover, informed consent was obtained from all respondents.

E. Data Analysis Procedure

Analysis followed the two-stage PLS-SEM approach. The measurement model was assessed first: indicator reliability (outer loadings ≥ 0.708, with 0.40–0.70 retained if AVE and composite reliability were satisfactory), internal consistency (Cronbach's α, ρA and ρC between 0.70 and 0.95), convergent validity (AVE ≥ 0.50), and discriminant validity using the Fornell–Larcker criterion, cross-loadings and the heterotrait–monotrait ratio (HTMT < 0.90). Indicator collinearity was checked using the variance inflation factor (VIF < 5). The structural model was then evaluated through path coefficients (β), bootstrapping with 5,000 subsamples for significance testing, the coefficient of determination (R^2), effect size (f^2) and predictive relevance (Q^2) using PLSpredict.

IV. RESULTS

A. Respondent Profile

The study includes 63 respondents, all from diverse academic and professional backgrounds. With respect to the respondent's highest qualification, 29% of the respondents possess a PhD, 24.7% of the respondents did an MD-PhD, 21.5% of the respondents had a DM degree, 19.4% of the respondents had an MD degree and the rest of them were had an equivalent specialized degree. This indicated participation from highly trained scientists and doctors. With regard to their primary place of work, 40.9% of the respondents were affiliated with research institutes, 23.7% of them with medical college/university, 22.6% of them had their private practice, and the rest 12.9% of respondents worked in hospitals. It was also observed that the respondents had a wide range of professional experience: 39.6% of respondents had 5-10 years of experience, 33% of them had 11-20 years of experience, 14.3% of them had more than 20 years of experience and 13.2% of them had less than 5 years of experience. These respondent characteristics indicate that most respondents were experienced professionals from academic, research, and clinical settings.

B. Measurement Model Assessment

Internal consistency and convergent validity. As shown in Table I, all three constructs exceeded the recommended thresholds. Cronbach's α ranged from 0.785 to 0.813, ρA from 0.814 to 0.842 and composite reliability (ρC) from 0.853 to 0.869, confirming internal consistency. The average variance extracted (AVE) ranged from 0.542 to 0.576, above the 0.50 benchmark, indicating that each construct explains more than half of the variance in its indicators.

Table I. Construct Reliability and Convergent Validity

Construct	Cronbach's α	ρA	ρC	AVE
AI-Driven Multi-Omics Integration (R1)	0.785	0.814	0.853	0.542
ML/DL Classification Capability (R2)	0.809	0.817	0.868	0.569
Explainability & Clinical Applicability (R3)	0.813	0.842	0.869	0.576

Indicator reliability and collinearity. Table II reports the outer loadings. Nine of the fifteen items loaded at or above 0.70, with the strongest indicators being items 14 (0.878), 4 (0.867) and 7 (0.822). Six items loaded between 0.568 and 0.697 (items 1, 5, 6, 11 and 12, and item 10 at 0.702). Following Hair et al. [cite], items between 0.40 and 0.70 were retained because their constructs still met the AVE and composite reliability criteria. Outer VIF values ranged from 1.486 to 3.307, well below 5, so indicator collinearity is not a concern.

Table II. Outer Loadings and Collinearity (VIF)

Construct	Item	Loading	VIF
AI-Driven Multi-Omics Integration (R1)	Item 1	0.609	1.598
	Item 2	0.771	1.855
	Item 3	0.752	1.615
	Item 4	0.867	2.145
	Item 5	0.654	1.501
ML/DL Classification Capability (R2)	Item 6	0.678	1.491
	Item 7	0.822	2.020

	Item 8	0.776	1.685
	Item 9	0.784	1.723
	Item 10	0.702	1.486
Explainability & Clinical Applicability (R3)	Item 11	0.568	1.585
	Item 12	0.697	1.801
	Item 13	0.819	2.098
	Item 14	0.878	3.307
	Item 15	0.793	2.326

Discriminant validity. Discriminant validity was assessed using three criteria (Tables III and IV). Under the Fornell–Larcker criterion, the square root of AVE for each construct (0.736, 0.754 and 0.759) was lower than its correlations with the other constructs (0.826–0.875). The HTMT ratios (1.012–1.058) also exceeded the conservative 0.90 threshold. The cross-loadings show that most items loaded highest on their own construct, but item 4 loaded almost as strongly on R2 (0.812) and R3 (0.814) as on R1 (0.867). Taken together, discriminant validity is not established: respondents did not clearly separate the three constructs. This is examined further in the Discussion and noted as a key limitation.

Table III. Fornell–Larcker Criterion

Construct	R1	R2	R3
R1	0.736		
R2	0.833	0.754	
R3	0.875	0.826	0.759

Table IV. Heterotrait–Monotrait Ratio (HTMT)

Construct	R1	R2	R3
R1	—		
R2	1.012	—	
R3	1.058	1.028	—

C. Structural Model Assessment

The structural model was evaluated through path coefficients, explanatory power, and effect sizes (Table V). AI-driven multi-omics integration showed a strong positive relationship with ML/DL classification capability ($\beta = 0.833$) and with explainability and clinical applicability ($\beta = 0.875$). The model explained 69.4% of the variance in classification capability ($R^2 = 0.694$; adjusted

R^2 = 0.691) and 76.5% of the variance in explainability and clinical applicability (R^2 = 0.765; adjusted R^2 = 0.762), indicating moderate-to-substantial explanatory power. The f^2 values of 2.272 and 3.253 far exceed the 0.35 benchmark for a large effect. Such unusually large effects are consistent with the high overlap between constructs noted above and should be interpreted with caution.

V. DISCUSSION

The study aims to understand how scientists and researchers perceive the advancement of having AI-driven multi-omic integration for drug discovery against cancer stem cells. The results indicated a very clear pattern on how this advancement is viewed: The respondents who rated AI based multi-omic integration to be effective also rated machine learning and deep learning classifiers to be efficient for drug discovery. Those respondents also expressed confidence in the value of explainable AI for clinical use.

Integration and classification (H1). The strong connection between integration and perceived classification capability (β = 0.833) contemplates a perspective of literature that single omic data is not efficient in identifying the plasticity and heterogeneity of CSCs (Chaudhary et al., 2017). Respondents believe that precise classification of CSCs is possible if the study is done with multi-omic deep learning models because they are confident that integrated data can outperform single-layer approaches in tumor subtyping and target identification (Subramanian et al., 2020).

Integration and explainability (H2). The even stronger link with explainability and clinical applicability (β = 0.875) suggests that respondents do not perceive predictive power to be adequate rather they believe that researchers need to reach to the understanding of why model flags a gene, pathway or protein to be able to trust AI-predicted targets in drug discovery. This supports the growing emphasis on XAI methods such as SHAP and attention mechanisms in precision oncology [cite]. A unified perception. The most notable finding is methodological. The lack of discriminant validity shows that respondents treated integration, classification and explainability as closely connected parts of one idea, the overall promise of AI in cancer research, rather than as separate capabilities. This may reflect genuine conceptual overlap, the early-career profile of many respondents, or item wording that did not separate the constructs sharply enough. It also explains the very large effect sizes and means the path results show a strong overall association rather than clearly distinct causal effects. Implications. For educators, the findings depict that they support and have enthusiasm for AI integrated methods but they also believe that education on the different roles of data integration, model selection and interpretability is much needed before starting to work on the integrated methods. For research teams, they suggest that interpretability should be built into multi-omics pipelines from the start rather than added later. In terms of precision oncology, researchers showed a positive signal for adoption but only if it is matched with rigorous empirical validation of AI models on real patient and cell-line data.

VI. LIMITATIONS

This study has a lot of limitations. The first is that it focuses on how different researchers feel about incorporating multi-omics and AI models rather than testing the actual performance of AI models and whether integrated models would work better than single omic approaches. Second is, discriminant validity was not established, which limits how far the three constructs can be interpreted separately. Third is, the purposive sample of 63 respondents, largely from respondents who have 5-10 years of professional experience, limits generalizability. Fourth is, the cross-sectional, self-reported design cannot establish causality and may be affected by common method bias. Finally, several indicators loaded below the preferred 0.70 threshold, suggesting that some items need refinement

VII. Conclusion and Future Scope

This study displays a proposed model of AI-driven multi-omics integration framework so that the exact therapeutics targets associated with cancer stem cells can be identified. Multi-omics data like genomic, transcriptomic, epigenomic, and proteomic information can help scientists identify the molecular characteristics that drive the stemness of cancer cells, the therapeutic resistance, and tumor heterogeneity. The study also proposes the idea of using machine learning and deep learning tools so that CSC can be classified accurately while explainable AI is incorporated to improve the interpretability of molecular predictions. The preliminary PLS-SEM results demonstrate satisfactory internal consistency and convergent validity for the three reported constructs. The structural model also exhibits R^2 values of 0.694 and 0.765. However, the HTMT and Fornell–Larcker assessments reveal concerns regarding discriminant validity, requiring further refinement of the measurement model.

The study proposes an integration of a lot of data and indicates that this kind of computational framework can be very beneficial for advancements in precision medicine since it allows for systematic candidate-target prioritization. Though the performance of the AI models and the biological ability of a computational framework to identify therapeutic targets still remains to be established. Though this study does emphasize the advancements in precision medicine and targeted drug discovery by integrating AI and multi-omic analysis, there is still a lack of rigorous empirical and experimental evaluation. Future scope in this field of research could have multi-omics data derived from standard public datasets like TCGA, CCLE TCGA, CCLE, GDSC and single-cell RNA-seq atlases. Research could focus on comparing integrated approaches and single-omics approaches for CSC target identification. An improvement in the survey could be having more clear and distinct items that are reevaluated from a larger sample of respondents from different countries. Having a larger sample size would provide data from respondents pursuing different professional careers like clinicians, bioinformaticians, and students that would help in making this study more reliable and accredited. Finally, therapeutic targets predicted by this computational framework could be tested through drug-repurposing screens and wet-lab experiments, and this would bridge the gap between perceived value of this research and its direct clinical impact.

REFERENCES

[1] Baysoy, A., Bai, Z., Satija, R., & Fan, R. (2023). The technological landscape and applications of single-cell multi-omics. Nature Reviews Molecular Cell Biology, 24(10), 695–713. https://doi.org/10.1038/s41580-023-00615-w

[2] Liu, Z., Xu, H., Weng, S., Ren, Y., & Han, X. (2022). Stemness refines the classification of colorectal cancer with stratified prognosis, Multi-Omics landscape, potential mechanisms, and treatment options. Frontiers in Immunology, 13, 828330. https://doi.org/10.3389/fimmu.2022.828330

[3] Catalano, M., D'Angelo, A., De Logu, F., Nassini, R., Generali, D., & Roviello, G. (2025). Navigating Cancer Complexity: Integrative Multi-Omics Methodologies for Clinical Insights. Clinical Medicine Insights Oncology, 19, 11795549251384582. https://doi.org/10.1177/11795549251384582

[4] Shrestha, P., Lee, D., Giri, A., Thi, T. L., & Yook, S. (2025). Genomic landscapes and therapeutic perspectives of cancer stem cells in the tumor microenvironment. Journal of Pharmaceutical Investigation, 56(4), 695–727. https://doi.org/10.1007/s40005-025-00794-y

[5] Huang, X., Huang, S., Reina, C., Šabanović, B., Roberto, M., Aicher, A., Tang, J., & Heeschen, C. (2025). Single-cell multi-omics and machine learning for dissecting stemness in cancer. Briefings in Bioinformatics, 26(5). https://doi.org/10.1093/bib/bbaf566

[6] Yuan, Q., Lu, X., Guo, H., Sun, J., Yang, M., Liu, Q., & Tong, M. (2023). Low-density lipoprotein receptor promotes crosstalk between cell stemness and tumor immune microenvironment in breast cancer: a large data-based multi-omics study. Journal of Translational Medicine, 21(1), 871. https://doi.org/10.1186/s12967-023-04699-y

[7] Vora, N., Shah, S., Patel, P., & Shah, M. (2025). Artificial intelligence and multi-omics in drug discovery: A deep learning-powered revolution. Cure & Care, 1(1), 100011. https://doi.org/10.1016/j.ccwv.2025.100011

[8] Sikder, T. R. T. R., Dash, S., Uddin, B., & Hossain, F. (2023). AI-Powered data Analytics and Multi-Omics integration for Next-Generation precision oncology and anticancer drug development. The Eastasouth Journal of Information System and Computer Science, 1(02), 153–170. https://doi.org/10.58812/esiscs.v1i02.838

[9] Mukherjee, A., Abraham, S., Singh, A., Balaji, S., & Mukunthan, K. S. (2024). From Data to Cure: A Comprehensive exploration of multi-omics data analysis for targeted therapies. Molecular Biotechnology, 67(4), 1269–1289. https://doi.org/10.1007/s12033-024-01133-6

[10] Albaradei, S. (2025). AI-Driven Approaches to Utilization of Multi-Omics Data for Personalized diagnosis and Treatment of Cancer: A Comprehensive Review. Computer Modeling in Engineering & Sciences, 145(3), 2937–2970. https://doi.org/10.32604/cmes.2025.072584

[11] Chakraborty, A., Gholap, A., Khuspe, P., Sundaram, G., Webster, T., Khalid, M., Haris, M., & Faiyazuddin, M. (2026). Advancing Drug Discovery with AI: Machine and Deep Learning Strategies for Target Identification and Precision Nanomedicine. International Journal of Nanomedicine, Volume 21, 1–62. https://doi.org/10.2147/ijn.s600651

[12] Wang, X., Xiong, D., Cui, S., Duan, B., Ding, G., Huang, Y., & Wang, Q. (2026). Artificial intelligence–enabled multi-omics biomarkers for immune checkpoint blockade: mechanisms, predictive modeling, and clinical translation. Frontiers in Immunology, 17, 1732079. https://doi.org/10.3389/fimmu.2026.1732079

[13] Albaradei, S. (2025b). AI-Driven Approaches to Utilization of Multi-Omics Data for Personalized diagnosis and Treatment of Cancer: A Comprehensive Review. Computer Modeling in Engineering & Sciences, 145(3), 2937–2970. https://doi.org/10.32604/cmes.2025.072584

[14] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M. Ringle, "When to use and how to report the results of PLS-SEM," European Business Review, vol. 31, no. 1, pp. 2–24, 2019, doi: 10.1108/EBR-11-2018-0203.

[15] J. Henseler, C. M. Ringle, and M. Sarstedt, "A new criterion for assessing discriminant validity in variance-based structural equation modeling," Journal of the Academy of Marketing Science, vol. 43, no. 1, pp. 115–135, 2015, doi: 10.1007/s11747-014-0403-8.

[16] C. Fornell and D. F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error," Journal of Marketing Research, vol. 18, no. 1, pp. 39–50, 1981, doi: 10.2307/3151312.

[17] Singh. R., Verghese. S. P., Bansal A. (2026). Silver Nanoparticles For Environmental Sustainability And Their Role In Biological Activities. International Journal of Engineering Sciences & Research Technology, 15(3), 1–8. ttps://doi.org/10.64149/j.ijesrt.15.3-1-8

[18] K. Chaudhary, O. B. Poirion, L. Lu, and L. X. Garmire, "Deep learning–based multi-omics integration robustly predicts survival in liver cancer," Clinical Cancer Research, vol. 24, no. 6, pp. 1248–1259, 2018, doi: 10.1158/1078-0432.CCR-17-0853.

[19] Nannuri S., Gantla H. R., Ananya D., Vennela A., Bhuvana B., Neha G. (2026). Automated Brain Tumor Segmentationin Mri Via U-Net-Based Deep Neural Networks. International Journal of Engineering Sciences & Research Technology, 15(4), 1–8. https://doi.org/10.64149/j.ijesrt.15.4.1-8

[20] I. Subramanian, S. Verma, S. Kumar, A. Jere, and K. Anamika, "Multi-omics data integration, interpretation, and its application," Bioinformatics and Biology Insights, vol. 14, 2020, doi: 10.1177/1177932219899051