International Journal of Public Health Science (IJPHS)
Vol. 14, No. 3, September 2025, pp. 1304~1312
ISSN: 2252-8806, DOI: 10.11591/ijphs.v14i3.25783

Computational paradigm for advancing lung cancer drug discovery

Ochin Sharma[1], Alwalid Bashier Gism Elseed Ahmed[2], Mudassir Khan[3], Ghantasala Gnana Sudha Pradeep[4], Pellakuri Vidyullatha[5], Mohammad Mazhar Nezami[6]
[1]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
[2]Department of Computer Science, College of Computer Science and Information Technology, University of Bisha, Bisha, Saudi Arabia
[3]Department of Computer Science, College of Computer Science, Applied College Tanumah, King Khalid University, Abha, Saudi Arabia
[4]Department of Computer Science and Engineering, Alliance University, Bengaluru, India
[5]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India
[6]Department of Computer Science, College of Computer Science and Information Technology, University of Bisha, Bisha, Saudi Arabia

Article Info

Article history:

Received Aug 22, 2024
Revised Jan 5, 2025
Accepted Mar 6, 2025

Keywords:

Bioinformatics
Employability
Industry
Medicine
Well- being

ABSTRACT

Lung cancer remainders one of the foremost causes of cancer-related impermanence worldwide. The availability of novel medicines for patients with lung cancer is restricted by the extremely lengthy timetables and high attrition rates of traditional drug discovery procedures. However, in silico drug discovery has emerged as a powerful and affordable way to identify potential treatments. This work offers well-structured paradigms for using virtual techniques to identify potential lung cancer treatments. The main concerns are virtual screening, target validation and identification, pharmacokinetic assessment, and molecular docking. The cost and time of drug development are reduced and a valuable platform for discovering novel drugs to treat lung cancer is produced by merging computational resources with proper methodologies. The current work explores the recent advancements, challenges, and possible future paths. Mann-Whitney U test says that the sampled data is different in distribution for molecular weight (MW), LogP, amount of H acceptors, and quantity of H donors for active and inactive molecules. Python tool has been utilized and identified that the CHEMBL4850929 ($C_{31}H_{31}F_2N_7O_4$) molecule is a potential drug. It has pIC50 7.61, Lipinski values in terms of MW 603.63, LogP 3.36, amount of H donors 1, quantity of H acceptors 10.

Corresponding Author:

Mudassir Khan
Department of Computer Science, College of Computer Science, Applied College Tanumah
King Khalid University
P.O. Box: 960 - Postal Code: 61421, Abha, Saudi Arabia
Email: mudassirkhan12@gmail.com

1. INTRODUCTION

Due to its high prevalence then terrible fatality rates, lung cancer poses a serious threat to global health. The process of developing effective medications for lung cancer is still time-consuming and resource-intensive, even with the notable advancements in cancer research and therapeutic interventions. The availability of novel medicines for patients with lung cancer is restricted by the extremely lengthy timetables and high attrition rates of traditional drug discovery procedures [1]–[5]. The convergence of in silico approaches, computational methodologies, and the Python programming language's adaptability has drawn increased attention in response to this urgent challenge. This convergence presents a ripe opportunity to transform the lung cancer medication discovery process. In silico drug, research unites data analysis,

simulations, and computer techniques to offer a novel and economical alternative. This research provides a comprehensive framework for lung cancer drug development that uses Python's computational capacity to discover, validate, and optimize viable lung cancer-fighting drug candidates more quickly [6]–[11].

This enterprise pursuits to bridge the space among computational procedures and experimental validation with the aid of using supplying a unified framework that leverages Python to speed up the drug discovery system. This technique includes, amongst different things, target selection, validation, molecular docking, virtual screening, and immoderate pharmacokinetic assessment [12]–[14]. The use of Python on this problematic system targets to boom the chance of coming across powerful lung cancer drug treatments even as decreasing the money and time lost on drug discovery. Researchers and pharmaceutical specialists can use this paper as a manual to help them make the maximum of Python and in silico drug improvement in the combat towards lung cancer and to enhance patient outcomes [15]–[19].

Since lung cancer has a chief effect on public health, it's far definitely important to create new and powerful drugs. Finding specialized treatment strategies is crucial seeing that lung cancer is available in a whole lot of forms, every with wonderful genetic and molecular characteristics [20]–[23]. Traditional drug improvement strategies may be highly priced and time-consuming, although they'll work in certain situations. On the alternative hand, the system of in silico drug discovery is quicker and much less highly priced, taking into consideration a methodical and data-driven look for capacity therapeutic candidates. This article outlines the manner for an in silico drug improvement framework and emphasizes the value of Python as a programming language. Python`s many tools, flexibility, and sturdy network make it a remarkable preference for imposing many computational components of drug studies [24]–[27]. Through the creation of a Python-based platform, this observe seeks to permit researchers, bioinformaticians, and computational biologists to accelerate the drug improvement manner and permit a faster translation of studies discoveries into beneficial medicines [28].

The value and time-ingesting nature of the traditional drug discovery technique are its drawbacks. The primary aim is to increase a drug discovery platform tailor-made to lung cancer, addressing this difficult trouble within silico techniques and the Python programming language. The number one task is growing an intensive in silico environment that encompasses all stages of the drug improvement technique, together with virtual screening, molecular docking, pharmacokinetic evaluation, target selection, and validation [29]–[31]. We desire to appoint Python-based libraries, computational tools, and machine-learning strategies to growth the performance of this procedure [32], [33]. This work aims to address the urgent need for a Python-driven, time-efficient, and cost-effective drug discovery framework to find auspicious therapeutic candidates for the dealing of lung cancer. The ultimate objective is to use Python's strengths in silico drug discovery to speed up the creation of new lung cancer treatments and enhance patient. Drug discovery is resource and time-consuming. Overall, the framework envisioned seeks to revolutionize drug discovery processes specifically for lung cancer while utilizing the power of Python and in silico methodologies [34]–[36].

The summary of the literature review has been summarized in Table 1. i) The researchers have observed that patients diagnosed with Lung cancer often face challenges in accessing proper treatment; ii) Furthermore, the study highlights the importance of early detection in improving patient outcomes, in general; iii) An integrated method for detecting lung cancer via CT scanning via optimization, deep learning, and IoT data transmission; iv) High-accuracy lung disease classification via logistic regression and advanced feature extraction techniques. In summary, the ongoing research efforts in Lung cancer are crucial for developing better treatment strategies and enhancing patient care. Based upon the literature review, a preclinical molecule is still to be identified, focusing on Lipinski values, excellent docking score, pIC values to act has a probable molecule into the cases for lung cancer.

Table 1. Summary of literature

S. No	Key findings and gaps	Methodology/approach	Citation
1	Identification of actionable targets is essential for lung cancer drug discovery. Genomic data, including the Cancer Genome Atlas (TCGA) and Genomic Data Commons (GDC), have been instrumental in identifying key driver mutations and pathways.	Computational biology, genomics, bioinformatics	[1]
2	Virtual screening is a critical step in silico drug discovery, enabling the efficient prioritization of compounds. Machine learning algorithms have been applied to enhance ligand-based virtual screening.	Virtual screening, machine learning	[2]
3	Targeted therapies for lung cancer, alike epidermal growth factor receptor (EGFR) and anaplastic lymphoma kinase (ALK) inhibitors, have been developed based on actionable targets identified using in silico approaches.	Target identification, drug design	[3]
4	Molecular docking imitations show a crucial role in predicting drug-target interactions. Advanced docking procedures, such as AutoDock Vina, have contributed to the discovery of potential lung cancer drug candidates.	Molecular docking, computational simulations	[4]
5	In silico drug discovery has the probable to rush the identification for innovative lung cancer drug candidates, plummeting time and cost. Advances in computational tools and data integration are shaping the future of the field.	In silico drug discovery, data integration	[5]

2. METHOD

Acidic mammalian chitinase (Chembl id: CHEMBL1293197) is identified as a target protein for further study to overcome lung cancer based upon literature review [1]–[5]. In the next step, a number of molecules have been listed out those can neutralize the target (Acidic mammalian chitinase protein). Founded onto the IC50 principles, all the molecules were considered into three categories: active, intermediate and inactive molecules.

The following methodological approach has been used:

i) Identification and validation of targets: Choosing and confirming a good therapeutic target is the first stage into the drug detection process. Potential targets for lung cancer could be particular proteins, receptors, or genetic abnormalities that are involved in the development and spread of the tumor. In silico strategies along with data mining, bioinformatics, and genomics evaluation are used to discover and validate those targets. This level determines whether or not the chosen goal is biologically substantial and contributes substantially to the sickness. Finding a success remedy alternative calls for an information of ways proteins, genetic information, and disease development interact.

ii) Pre-screening: Chemical compounds in massive molecular libraries are looked after and ranked the use of a computational technique called virtual screening. This segment of the lung cancer drug discovery procedure consists of searching through chemical databases for capability treatment alternatives that would engage with the selected goal. Virtual screening strategies based on shape and molecular modeling are used to evaluate a compound`s binding affinities with target proteins. The accuracy of molecular interactions and predictive modeling are critical to this procedure's success.

iii) Pharmacokinetic evaluation: Understanding the pharmacokinetic characteristics of therapeutic applicants, inclusive of absorption, distribution, metabolism, and excretion (ADME), has a essential phase in the drug improvement progression. When assessing the pharmacokinetic residences of viable therapeutic drugs, predictive strategies and in silico models are essential. The efficacy of the scientific trials may be expected via way of means of deciding on drug applicants with favorable pharmacokinetic properties.

iv) Docking of molecules: Molecular docking is a important level in the in silico drug discovery process. Modeling the interplay among the drug candidate and the target protein was essential to are expecting binding affinities, mechanisms of binding, and capacity drug-protein interactions. Molecular docking investigations are conducted using various techniques and software tools, to ascertain the most suitable drug candidates for further development. Validation of the interactions between the drug and target proteins is crucial for the success of therapeutic interventions.

v) Current progress and difficulties: This paper reviews the current progress in silico lung cancer drug discovery, focusing on the latest progressions into artificial intelligence then machine learning for predictive modelling. Additionally, the study of protein-ligand interactions using molecular dynamics simulations provides valuable acumens into the machineries of deed of potential drugs. The integration of structural biology data into drug design has accelerated the drug discovery process for lung cancer, despite challenges in obtaining accurate predictive models. Overall, the integration of advanced computational methods with biological insights holds promise for the continued advancement of lung cancer drug discovery efforts. The confirmation of computer predictions through experimental research, the requirement for precise and trustworthy datasets, and the ethical issues surrounding the use of patient data are some of the challenges.

vi) Future expectations: Lung cancer in silico drug discovery has a lot of potential, and more research should improve the process's accuracy and efficiency. Prospects for the future include integrating multi-omics data to find novel targets and therapeutic candidates, developing more precise predictive models, and using personalized medicine techniques. Additionally, in silico drug discovery's computational capabilities will keep getting better because of developments in high-performance computing and cloud-based resources.

3. RESULTS AND DISCUSSION

CheMBL database is a web repository. An application programming interface (API) of this web repository has been used with Python to extract proteins associated in the lung cancer. Acidic mammalian chitinase, CheMBL id: CHEMBL1293197 is identified as a target protein for further study to overcome lung cancer. In the next step, several molecules have been listed that can neutralize the Acidic mammalian chitinase protein. Based on the half maximal inhibitory concentration (IC50) standards, all the molecules were characterized into three categories: active, intermediate and inactive molecules.

Table 2, shows the results of active vs inactive molecules using the Mann-Whitney U test in terms of comparing data distribution of LipinSki measures: logarithm of the partition coefficient (Logp), molecular weight (MW), H-acceptors, H-donors and negative logarithm of (pIC50). Mann-Whitney U test has the independent sample testing alternate to the non-parametric t-test. The non-parametric test comparations two sample means after the equivalent population which limits whether or not has the two sample means were

equivalent. The results depict that active and inactive molecules have a different distribution in terms of Lipinski measure and studying active molecules only is worthwhile during the research.

Figure 1 is helpful to observe the whole dataset in terms bioactivity of molecules vs LipinSki values. In Figure 1(a), shows those molecules seems more active which is more than 8 H donors. Mostly inactive molecules have H donors values between 2 to 3 Hydrogen bond donors. Figure 1(b), shows bar chart for bioactivity vs frequency of molecules. It can be observed that active molecules have higher frequency. Figure 1(c) shows that bioactive molecules are having higher pIC50 values. Figure 1(d), shows that molecules are having higher LogP values are more bioactive. Figure 1(e), shows that molecules are having higher LogP values are more bioactive. Ideally, a molecule should accept more than 5 hydrogen atoms. Figure 1(f), shows molecules under observation has hydrogen atoms between 5 to 6 wheather bioactivity class is active or inactive.

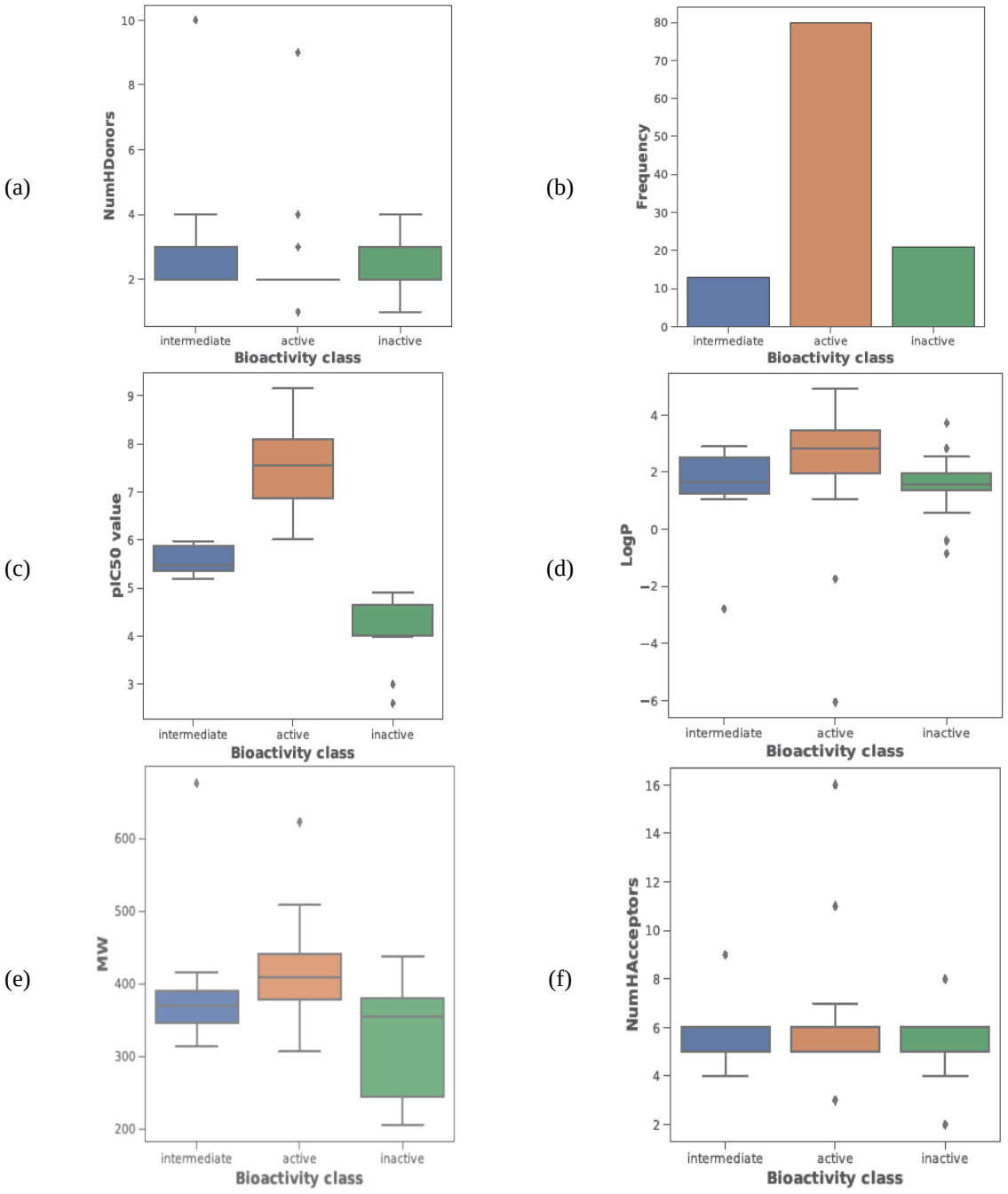

Figure 1. Dataset in terms bioactivity of molecules vs LipinSki values: (a) bioactivity vs H donors, (b) bioactivity vs frequency of molecules, (c) bioactivity vs molecules pIC50, (d) bioactivity vs molecules LogP, (e) bioactivity vs molecular weight, and (f) bioactivity vs H acceptors

Table 3, represents the LipinSki values which is to identify in terms of a better drug molecule. The purpose of the Lipinski measure is to explore drug absorption and extraction from the body. In this measure, hydrogen bond donors must be less than 5 donors must be less than 10, octanol-water partition should not exceed 5. Based on these values, five molecules have been listed. Further, based on the pIC50 value, the CHEMBL4850929 molecule is considered a drug candidate. Let's see its docking score also. Figure 2, represents the molecular formula for CHEMBL4850929 this molecule is in a preclinical state and has not been an established drug till now. Further candidates for clinical trials can be considered for this.

Figure 3, represents the docking score of CHEMBL4850929 is -10.9 which is considered a very good score. Docking is an essential technique in drug development as it shows how the proposed drug molecule can be absorbed and how easily the residue of the molecule will be excreted from the patient's body. Figure 4, represents the docking diagram of the CHEMBL4850929 molecule, which shows that the proposed drug molecule can be well absorbed by a patient.

Table 2. Mann-Whitney U test with active vs inactive molecules

Descriptor	Statistics	Log P	Alpha	Interpretation
LogP	808.5	0.001417	0.05	Dissimilar distribution (reject H0)
Molecular weight (MW)	798	0.002105	0.05	Different distribution (reject H0)
NumHAcceptors	653	0.10833	0.05	Same distribution (fail to reject H0)
NumHDonors	377.5	0.045048	0.05	Different distribution (reject H0)
pIC50	1040	8.60E-09	0.05	Dissimilar distribution (reject H0)

Table 3. Active molecules identified based upon highest values of pIC50

Molecule_chembl_id	Canonical_smiles	Bioactivity	MW	LogP	H donors	H acceptors	pIC 50
CHEMBL4850929	CC(C)n1cc(C(=O)Nc2ccc(Oc3ncnc4c3CCN(C3CN(C)C3)C4)c(F)c2)c(=O)n(-c2ccc(F)cc2)c1=O	active	603	3.36	1	10	7.61
CHEMBL4878216	CC(C)n1cc(C(=O)Nc2ccc(Oc3ncnc4c3CN(C(=O)C3CC3)C4)c(F)c2)c(=O)n(-c2ccc(F)cc2)c1=O	active	588	3.94	1	9	7.46
CHEMBL4860541	CN1CCC(CN2CCc3c(ncnc3Oc3ccc(NC(=O)c4cn(C5CCCC5)c(=O)n(-c5ccc(F)cc5)c4=O)cc3F)C2)CC1	active	671	4.92	1	10	7.46
CHEMBL4865150	CC(C)n1cc(C(=O)Nc2ccc(Oc3ncnc4c3CN(C(=O)CC3CCN(C)CC3)C4)c(F)c2)c(=O)n(-c2ccc(F)cc2)c1=O	active	659	3.36	1	10	7.40
CHEMBL4858396	CN1CCC(CN2CCc3c(ncnc3Oc3ccc(NC(=O)c4cn(CC5CC5)c(=O)n(-c5ccc(F)cc5)c4=O)cc3F)C2)CC1	active	657	3.94	1	10	7.29

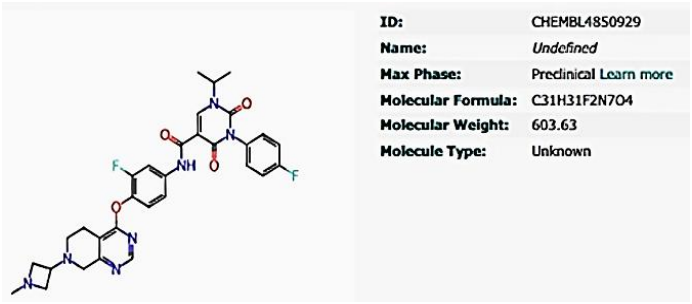

Figure 2. CHEMBL4850929 (C31H31F2N7O4)

Figure 3. Docking results of CHEMBL4850929 (C31H31F2N7O4)

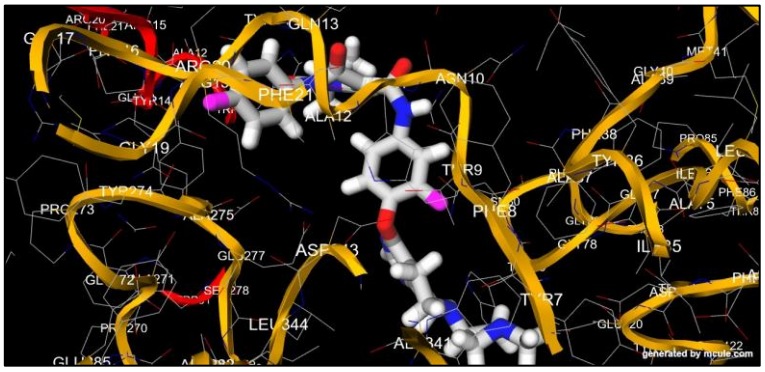

Figure 4. Docking results of CHEMBL4850929 (C31H31F2N7O4) for a score of -10.9

4. CONCLUSION

The modern study presents an in-silico drug discovery framework that provides a scientific to identifying potential novel medicines for the treatment of lung cancer. By the usage of computational equipment and related processes, this method reduces expenses and expedites the drug discovery process. As the sector advances, the improvement of modern lung cancer drugs may be notably impacted through in silico drug discovery, which could finally enhance patient outcomes. Based at the pIC50 and Lipinski values, the have a look at tested that CHEMBL4850929 (C31H31F2N7O4) is a possible treatment choice for lung cancer. Additionally, docking scores, which are -10.9, desire C31H31F2N7O4. Additionally, medical trials are being carried out at the chemical.

ACKNOWLEDGMENTS

The University of Bisha`s Deanship of Graduate Studies and Scientific Research supplied investment for these works via the Fast-Track Research Support Program, for which the authors are grateful.

FUNDING INFORMATION

The authors declare that there is no funding to be acknowledged.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ochin Sharma	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Alwalid Bashier Gism Elseed Ahmed		✓				✓		✓	✓	✓	✓	✓		
Mudassir Khan	✓		✓	✓			✓			✓	✓		✓	✓
Ghantasala Gnana Sudha Pradeep	✓	✓		✓		✓			✓		✓			
Pellakuri Vidyullatha					✓		✓			✓		✓		✓
Mohammad Mazhar Nezami	✓			✓					✓	✓				✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ETHICAL APPROVAL

The conducted research is not related to either human or animal use.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, [MK], the data are not publicly available due to privacy or ethical restrictions.

REFERENCES

[1] Reem Nooreldeen and Horacio Bach, "Current and future development in lung cancer diagnosis," *International Journal of Molecular Sciences*, vol. 22, no. 8661, p. 1, 2021.

[2] C.-H. Liu *et al.*, "DriverDBv4: a multi-omics integration database for cancer driver gene research," *Nucleic Acids Research*, vol. 52, no. D1, pp. D1246–D1252, Jan. 2024, doi: 10.1093/nar/gkad1060.

[3] T. A. Soares, A. Nunes-Alves, A. Mazzolari, F. Ruggiu, G. W. Wei, and K. Merz, "The (Re)-Evolution of quantitative structure-activity relationship (QSAR) studies propelled by the surge of machine learning methods," *Journal of Chemical Information and Modeling*, vol. 62, no. 22, pp. 5317–5320, 2022, doi: 10.1021/acs.jcim.2c01422.

[4] A. Leiter, R. R. Veluswamy, and J. P. Wisnivesky, "The global burden of lung cancer: current status and future trends," *Nature Reviews Clinical Oncology*, vol. 20, no. 9, pp. 624–639, 2023, doi: 10.1038/s41571-023-00798-3.

[5] C. Kirubhanand *et al.*, "Targeting potential receptor molecules in non-small cell lung cancer (NSCLC) using in silico approaches," *Frontiers in Molecular Biosciences*, vol. 10, 2023, doi: 10.3389/fmolb.2023.1124563.

[6] D. K. Chebanov, V. A. Misyurin, and I. Z. Shubina, "An algorithm for drug discovery based on deep learning with an example of developing a drug for the treatment of lung cancer," *Frontiers in Bioinformatics*, vol. 3, 2023, doi: 10.3389/fbinf.2023.1225149.

[7] P. Du, R. Fan, N. Zhang, C. Wu, and Y. Zhang, "Advances in integrated multi-omics analysis for drug-target identification," *Biomolecules*, vol. 14, no. 6, 2024, doi: 10.3390/biom14060692.

[8] S. Pandiyan and L. Wang, "A comprehensive review on recent approaches for cancer drug discovery associated with artificial intelligence," *Computers in Biology and Medicine*, vol. 150, 2022, doi: 10.1016/j.compbiomed.2022.106140.

[9] N. A. Murugan, A. Podobas, D. Gadioli, E. Vitali, G. Palermo, and S. Markidis, "A review on parallel virtual screening softwares for high-performance computers," *Pharmaceuticals*, vol. 15, no. 1, 2022, doi: 10.3390/ph15010063.

[10] Q. Li and C. Kang, "Mechanisms of action for small molecules revealed by structural biology in drug discovery," *International Journal of Molecular Sciences*, vol. 21, no. 15, pp. 1–18, 2020, doi: 10.3390/ijms21155262.

[11] G. Choi, D. Kim, and J. Oh, "AI-based drug discovery of TKIs targeting L858R/T790M/C797S-Mutant EGFR in Non-small cell lung cancer," *Frontiers in Pharmacology*, vol. 12, 2021, doi: 10.3389/fphar.2021.660313.

[12] M. Bittlinger, S. Bicer, J. Peppercorn, and J. Kimmelman, "Ethical considerations for phase I trials in oncology," *Journal of Clinical Oncology*, vol. 40, no. 30, pp. 3474–3488, 2022, doi: 10.1200/JCO.21.02125.

[13] N. Faruqui, M. A. Yousuf, F. A. Kateb, M. Abdul Hamid, and M. M. Monowar, "Healthcare as a service (HAAS): CNN-based cloud computing model for ubiquitous access to lung cancer diagnosis," *Heliyon*, vol. 9, no. 11, 2023, doi: 10.1016/j.heliyon.2023.e21520.

[14] S. Rezayi, S. R Niakan Kalhori, and S. Saeedi, "Effectiveness of artificial intelligence for personalized medicine in neoplasms: A systematic review," *BioMed Research International*, vol. 2022, 2022, doi: 10.1155/2022/7842566.

[15] R. Paudyal *et al.*, "Artificial intelligence in CT and MR imaging for oncological applications," *Cancers*, vol. 15, no. 9, 2023, doi: 10.3390/cancers15092573.

[16] M. Kumari, "Computational approaches streamlining molecular modelling and drug designing," *Research Spectra*, vol. 4, no. 2, pp. 76–125, 2023.

[17] W. Chen, S. Zhu, Y. Zhang, J. Xiao, and D. Tian, "Identification of key candidate tumor biomarkers in non-small-cell lung cancer by in silico analysis," *Oncology Letters*, vol. 19, no. 1, pp. 1008–1016, 2020, doi: 10.3892/ol.2019.11169.

[18] S. Kour *et al.*, "Artificial intelligence and nanotechnology for cervical cancer treatment: Current status and future perspectives," *Journal of Drug Delivery Science and Technology*, vol. 83, 2023, doi: 10.1016/j.jddst.2023.104392.

[19] M. Tamminga, T. J. N. Hiltermann, E. Schuuring, W. Timens, R. S. N. Fehrmann, and H. J. M. Groen, "Immune microenvironment composition in non-small cell lung cancer and its association with survival," *Clinical and Translational Immunology*, vol. 9, no. 6, 2020, doi: 10.1002/cti2.1142.

[20] M. K. Park *et al.*, "Deep-learning algorithm and concomitant biomarker identification for NSCLC prediction using multi-omics data integration," *Biomolecules*, vol. 12, no. 12, 2022, doi: 10.3390/biom12121839.

[21] L. Chang, J. Wu, N. Moustafa, A. K. Bashir, and K. Yu, "AI-driven synthetic biology for non-small cell lung cancer drug effectiveness-cost analysis in intelligent assisted medical systems," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 10, pp. 5055–5066, 2022, doi: 10.1109/JBHI.2021.3133455.

[22] M. A. U. Khalid, Y. S. Kim, M. Ali, B. G. Lee, Y. J. Cho, and K. H. Choi, "A lung cancer-on-chip platform with integrated biosensors for physiological monitoring and toxicity assessment," *Biochemical Engineering Journal*, vol. 155, p. 107469, 2020, doi: 10.1016/j.bej.2019.107469.

[23] A. Olğaç, A. Türe, S. Olğaç, and S. Möller, "Cloud-based high throughput virtual screening in novel drug discovery," *Lecture Notes in Computer Science* 2019, pp. 250–278. doi: 10.1007/978-3-030-16272-6_9.

[24] M. Heidelberg, A. Kelman, J. Hopkins, and M. E. Allen, "The evolution of data ethics in clinical research and drug development," *Ethics, Medicine and Public Health*, vol. 14, p. 100517, 2020, doi: 10.1016/j.jemep.2020.100517.

[25] M. J.D., R. J.A., and G. A.F., "Focus on lung cancer," *Cancer Cell*, vol. 1, no. 1, pp. 49–52, 2002.

[26] M.-H. Tao, "Epidemiology of lung cancer," in *Lung Cancer and Imaging*, IOP Publishing, 2019, pp. 4-1-4–15. doi: 10.1088/978-0-7503-2540-0ch4.

[27] S. K. Thakur, D. P. Singh, and J. Choudhary, "Lung cancer identification: a review on detection and classification," *Cancer and Metastasis Reviews*, vol. 39, no. 3, pp. 989–998, 2020, doi: 10.1007/s10555-020-09901-x.

[28] S. C. Spiro and G. A. Silvestri, "One hundred years of lung cancer," *American Journal of Respiratory and Critical Care Medicine*, vol. 172, no. 5, pp. 523–529, 2005, doi: 10.1164/rccm.200504-531OE.

[29] M. K. Thirunavukkarasu, U. Suriya, T. Rungrotmongkol, and R. Karuppasamy, "In silico screening of available drugs targeting non-small cell lung cancer targets: A drug repurposing approach," *Pharmaceutics*, vol. 14, no. 1, p. 59, 2022, doi: 10.3390/pharmaceutics14010059.

[30] T. Muthu Kumar and K. Ramanathan, "Revealing the potency of natural compounds as dual kinase blockers by structure-based modelling and extensive MD simulation strategy," *Journal of Computational Biophysics and Chemistry*, vol. 21, no. 5, pp. 515–528, 2022, doi: 10.1142/S273741652250017X.
[31] P. Ramesh and S. Veerappapillai, "Designing novel compounds for the treatment and management of RET-positive non-small cell lung cancer—fragment based drug design strategy," *Molecules*, vol. 27, no. 5, p. 1590, 2022, doi: 10.3390/molecules27051590.
[32] P. Ramesh, W. H. Shin, and S. Veerappapillai, "Discovery of a potent candidate for ret-specific non-small-cell lung cancer—a combined in silico and in vitro strategy," *Pharmaceutics*, vol. 13, no. 11, p. 1775, 2021, doi: 10.3390/pharmaceutics13111775.
[33] O. O. Elekofehinti *et al.*, "Discovery of potential epidermal growth factor receptor inhibitors from black pepper for the treatment of lung cancer: an in-silico approach," *In Silico Pharmacology*, vol. 12, no. 1, p. 28, 2024, doi: 10.1007/s40203-024-00197-1.
[34] J. E. Faya Castillo, R. J. Zapata Dongo, P. A. Wong Chero, and S. F. Infante Varillas, "Mitoxantrone and abacavir: An ALK protein-targeted in silico proposal for the treatment of non-small cell lung cancer," *PloS one*, vol. 19, no. 2, p. e0295966, 2024, doi: 10.1371/journal.pone.0295966.
[35] S. Karimullah, M. Khan, F. Shaik, B. Alabduallah, and A. Almjally, "An integrated method for detecting lung cancer via CT scanning via optimization, deep learning, and IoT data transmission," *Frontiers in Oncology*, vol. 14, Oct. 2024, doi: 10.3389/fonc.2024.1435041.
[36] S. Kaur *et al.*, "High-accuracy lung disease classification via logistic regression and advanced feature extraction techniques," *Egyptian Informatics Journal*, vol. 29, p. 100596, Mar. 2025, doi: 10.1016/j.eij.2024.100596.

BIOGRAPHIES OF AUTHORS

Ochin Sharma is a Full-time Professor in the Department of CSE, Chitkara University, Punjab. He earned an M.Tech. (CSE) from Kurukshetra University, India and Ph.D. from Banasthali University, India. His areas of interest for research are machine learning, data analytics, and bio-informatics. He has taught various subjects at Banasthali University, Manav Rachna International Institute of Research & Studies, Chitkara University. He can be contacted at email: ochin.sharma@gmail.com.

Alwalid Bashier Gism Elseed Ahmed his bachelor Bachelor's degree from Al-Neelain University of Science and Technology in 2002. He received a Master's degree from Al-Neelain University-College of Computer Science and Information Technology in 2008 and a Ph.D. degree from Al-Neelain University-College of Graduate Studies in 2015. His major research areas include E-government applications, artificial intelligence, and big data. He can be contacted at email: alwlidbasheer@ub.edu.sa.

Mudassir Khan is a distinguished academician, researcher, author, and consultant with a robust background in computer science and data analytics. He earned his Master of Computer Applications (MCA) from Gautam Budh Technical University, India, and subsequently completed his Ph.D. in Big Data Analytics using Deep Learning from Noida International University, India. Currently, he is pursuing a postdoctoral fellowship in the Faculty of Computing Informatics at Multimedia University, Malaysia. He is Co-Supervisor for Postgraduate Studies in the Centre of Postgraduate Studies, Department of Computer Science, Lincoln University College, Malaysia. With over 13+ years of teaching and research experience, Dr. Khan has made significant contributions to the field while serving as an Assistant Professor in the Department of Computer Science at the College of Computer Science, Applied College, Tanumah, King Khalid University, Kingdom of Saudi Arabia. His expertise encompasses a broad spectrum of subjects, including big data, deep learning, machine learning, data science, the internet of things (IoT), and artificial intelligence, medical imaging particularly in the context of healthcare. Dr. Khan is a prolific author, having published over 55+ research articles in prestigious international referred journals, including those indexed by SCI, SCIE, and SCOPUS. His work has also been presented at various international conferences, with publications indexed in Springer, IEEE Xplore, and ACM. Dr. Khan has edited numerous books published by esteemed publishers such as Springer, Wiley, Bentham Science, Nova Publishers, and Elsevier. In addition to his articles, he has authored four books in the field of computer science. He can be contacted at email: mudassirkhan12@gmail.com.

Ghantasala Gnana Sudha Pradeep is currently a Professor at the Department of Computer Science and Engineering, Alliance University, Bengaluru, India. He has 17+ years of academic experience. He completed his Post-Doctoral at Industrial University of Ho Chi Minh City, Vietnam. He received a Ph.D. degree in computer science and engineering from Shri Venkateshwara University, India, in 2018, an M.Tech. degree in computer science and engineering from Acharya Nagarjuna University, India, in 2009, a B.Tech. degree in information technology from Jawaharlal Nehru Technological University Hyderabad, India, in 2006. He received more than six awards for his academic excellence. Under his supervision Ph.D. scholars are working and he is the expert evaluation panel member for various reputed universities. He has published various papers in reputed journals SCI; Scopus-indexed journals, UGC journals, conferences, books, book chapters, and around 45 national and international granted and published patents. He is also the editor and reviewer in various journals. His research interests include machine learning, deep learning, healthcare applications, and software engineering applications. He can be contacted at email: ggspradeep@gmail.com.

Pellakuri Vidyullatha completed her postdoctoral fellowship in 2023 at the Industrial University of Ho Chi Minh City, Vietnam, focusing on "An Efficient Data Analytics and Optimized Algorithm for Enhancing the Performance of Image Processing Using Artificial Intelligence." Prior to this, she earned her Master of Technology in computer science from JNTU-Hyderabad and pursued her Ph.D. in the Department of Computer Science & Engineering at KL University, Guntur, AP. Currently serving as an Associate Professor in the Department of Computer Science & Engineering at Koneru Lakshmaiah Education and Foundation in Guntur, Andhra Pradesh, she has received numerous accolades including awards for Best Researcher, Outstanding Researcher, and recognition as an Inspiring Woman. Her contributions extend to skill excellence, delivering guest lectures on deep learning and machine learning models. She holds global certifications from Google TensorFlow and Google Data Engineering. Her expertise spans artificial intelligence, machine learning, and deep learning techniques, with a focus on image analysis, text analysis, big data analytics, and the internet of things (IoT). She has published extensively in over 50 Scopus indexed journals and international conferences. Membership in prestigious organizations such as IAENG, IACSIT, and SDIWC further solidify her standing in the academic community. Additionally, she has secured both national and international patents, with grants in the realms of deep learning and IoT, underscoring her impact on cutting-edge research and innovation in her field. She can be contacted at email: latha22pellakuri@gmail.com.

Mohammad Mazhar Nezami is currently serving as a Lecturer in the Department of Computer Science at the College of Computing and Information Technology, University of Bisha. With over 14 years of enriching academic experience, he has established himself as a passionate educator and an active contributor to the field of computer science. He earned his bachelor's degree in information technology and master's degree in computer science and applications from Aligarh Muslim University, India, in 2006 and 2009, respectively. His commitment to advancing knowledge is reflected in his ongoing pursuit of a Ph.D. in Computer Science. His research interests include cloud computing, workflow scheduling, meta-heuristics optimization, and multi-objective optimization. His expertise extends to developing innovative optimization algorithms aimed at enhancing computational efficiency in cloud-based systems. He can be contacted at email: ndhami@ub.edu.sa.