Peer Review File

Article Information: https://dx.doi.org/10.21037/tlcr-23-350

Reviewer A

As they know, the main limitation of AI introduction in clinical practice is its difficulty to correlate it with clinically relevant specific variables.

The authors have included several variables in their model, but the clinical stratification of these variables remain unclear to me.

To add novelty to the authors I would encourage to take into account other variables as Stage, quality of surgery, quality of systemic treatment, etc.

There are other issues:

1. In the introduction, I would eliminate the epidemiology and the treatments part, as I think it is repeated and not relevant at that point.

2. In the discussion, the authors fail to introduce clinically relevant aspects and they make controversial statements (as the preferred surgery was the less extensive resection). I would recommend revision.

3. Comparing the outcomes of their model to clinicians results would add value to the manuscript.

4. Lack of definition of tumor recurrence in the manuscript.

Thank you for your *comments* on our article. With reference to the reviewer's comments:

The study includes cases of patients with confirmed lung cancer diagnosis. Detailed socioeconomic and treatment data are presented in Table 1. Due to the multitude of chemotherapy used, depending on the histopathological type of cancer, the patient's clinical condition, comorbidities, and the multitude of regimens, the work does not include a precise division into the type of chemotherapeutic agent. However, we have taken into account a type of treatment in diviation to their mechanism - a division into chemotherapy, molecularly targeted treatment, and drugs that stimulate the immune system. The problem is described in the limitation section (line 403-405). In addition, in the Methods section, data on the included division have been supplemented.

Ad.1: The introduction of the work was modified, the part concerning epidemiology and treatment was removed.

Ad.2.: Part of the discussion was modified according to the reviewer's recommendation

Ad.3.: In accordance with the reviewer's comment, as well as the reviewer's comment No. 2, information on the comparison of the model with the data from the literature was added in the paper, as well as a short review of the literature from the last 5 years regarding the treatment methods used. (line 317-339, Table 5)

Ad.4.: Tumor recurrent definition added (line 172-175)

"Recurrence was defined as a local, regional or distant recurrence based on the North American Association of Central Registries (NAACCR), ACoS – Commission on Cancer and SEER/NCI in accordance with the Standards for Oncology Registry Entry (STORE) guidelines"

Reference:

19. American College of Surgeons Commission on Cancer. STandards for Oncology Registry Entry (STORE) v. 1.0 [Internet]. 2018 [cited 2023 Jul 23]. Available from: https://www.facs.org/-/media/files/quality-programs/cancer/ncdb/store_manual_2018.ashx

Reviewer B

Comment 1. In the Introduction, specifically in lines 110-122, starting with "Artificial Intelligence (AI) techniques…" Authors introduce the AI and explain the importance of Artificial Neural Networks (ANNs). However, I think Authors could expand more this section stressing more the role played by these ANNs. Indeed, through the years, various medical studies, including diseases of brain, breast, lung, melanoma etc. have used ANNs, in particular the Convolutional Neural Networks (CNNs), reaching good performances. Therefore, Authors could include these works as references:

• R. Bellotti et al. "Topographic classification of EEG patterns in Huntington's disease", Neurol. Clin. Neuro- physiol. 2004 Nov 30; 2004:37.
• M.C. Comes et al., "Early Prediction of Breast Cancer Recurrence for Patients Treated with Neoadjuvant Chemotherapy: A Transfer Learning Approach on DCE-MRI", Cancers 13, 2298 (2021).
• M. El Adoui et al., "Multi-input deep learning architecture for predicting breast tumor response to chemotherapy using quantitative MR images", Int. J. Comp. Ass. Rad. Surg. 15, 1491-1500 (2020).
• M.C. Comes et al., "A deep learning model based on whole slide images to predict disease-free survival in cutaneous melanoma patients", Sci. Rep. 12, 20366 (2022).
• T.J. Brinker et al., "Deep learning approach to predict sentinel lymph node status directly from routine histology of primary melanoma tumours", Eur. J. Cancers 154, 227-234 (2021).
• S. Bove et al., "A CT-based transfer learning approach to predict NSCLC recurrence: The added-value of peritumoral region", PLoS ONE 18(5), e0285188 (2023).

Answer:

In accordance with the reviewer's comment, information about artificial intelligence usage was added in the introduction part (line 112-122), and reference was made to the literature data.

References:

9. Bellotti R, De Carlo F, Massafra R, de Tommaso M, Sciruicchio V. Topographic classification of EEG patterns in Huntington's disease. Neurol Clin Neurophysiol. 2004 Nov 30;2004:37.

10. Comes MC, Forgia D La, Didonna V, Fanizzi A, Giotta F, Latorre A, et al. Early Prediction of Breast Cancer Recurrence for Patients Treated with Neoadjuvant Chemotherapy: A Transfer Learning Approach on DCE-MRIs. Cancers (Basel) [Internet]. 2021 May 11 [cited 2023 Jul 19];13(10)(2298). Available from: https://pubmed.ncbi.nlm.nih.gov/34064923/

11. El Adoui M, Drisis S, Benjelloun M. Multi-input deep learning architecture for predicting breast tumor response to chemotherapy using quantitative MR images. Int J Comput Assist Radiol Surg [Internet]. 2020 Sep 1 [cited 2023 Jul 19];15(9):1491–500. Available from: https://pubmed.ncbi.nlm.nih.gov/32556920/

12. Bove S, Fanizzi A, Fadda F, Comes MC, Catino A, Cirillo A, et al. A CT-based transfer learning approach to predict NSCLC recurrence: The added-value of peritumoral region. PLoS One [Internet]. 2023 May 2 [cited 2023 Jul 19];18(5). Available from: https://pubmed.ncbi.nlm.nih.gov/37130116/

13. Comes MC, Fucci L, Mele F, Bove S, Cristofaro C, De Risi I, et al. A deep learning model based on whole slide images to predict disease-free survival in cutaneous melanoma patients. Sci Rep [Internet]. 2022 Nov 27 [cited 2023 Jul 19];12(20366):1–10. Available from: https://www.nature.com/articles/s41598-022-24315-1

14. Brinker TJ, Kiehl L, Schmitt M, Jutzi TB, Krieghoff-Henning EI, Krahl D, et al. Deep learning approach to predict sentinel lymph node status directly from routine histology of primary melanoma tumours. Eur J Cancer [Internet]. 2021 Sep 1 [cited 2023 Jul 19];154:227–34. Available from: https://pubmed.ncbi.nlm.nih.gov/34298373/

Comment 2. In the section Statistical Analysis, the database is introduced with the adopted clinical features with reference to Table 1. I'd suggest the Authors to be uniform in the choice of symbols , and . for thousand and decimals. In the Tables and also frequently in the text , and . seem to be used alternatively and this is quite confusing. I'd also suggest the Authors to be careful in the choice of number format for the metrics (e.g., whether decimal or in percentage).

Answer:

According to the reviewer's suggestion, the manuscript was carefully revised focusing on mentioned issues. All the symbols were uniformed and mistakes were corrected.

The number format is dependent on the variable type. The quantitative variables are presented as decimals with SD and qualitative variables as percentages of the whole group.

Comment 3. Figures 1 and 2 show small writings and they are not so easy to read. I'd suggest a bigger font.

Answer:

The writings on the figures were enlarged to make it easier to read.

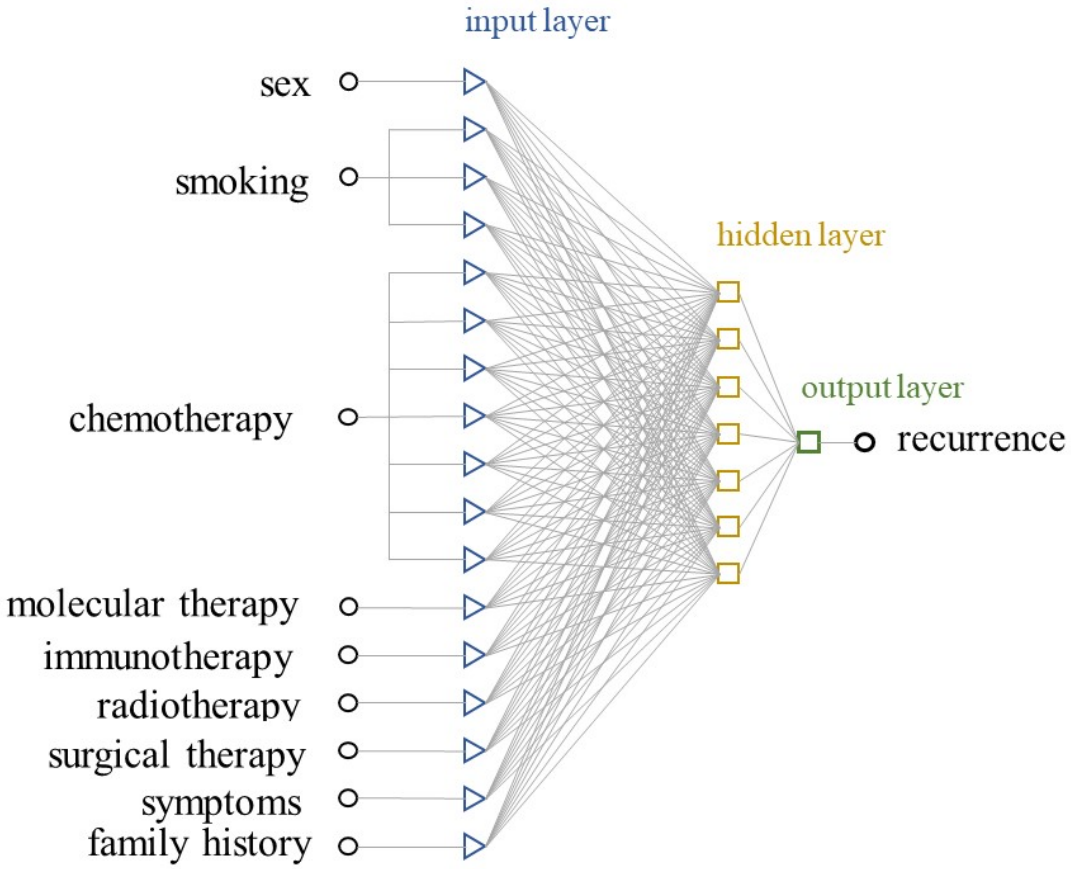

Figure 1. Graph of selected ANN model - MLP 9:17-7-1:1.

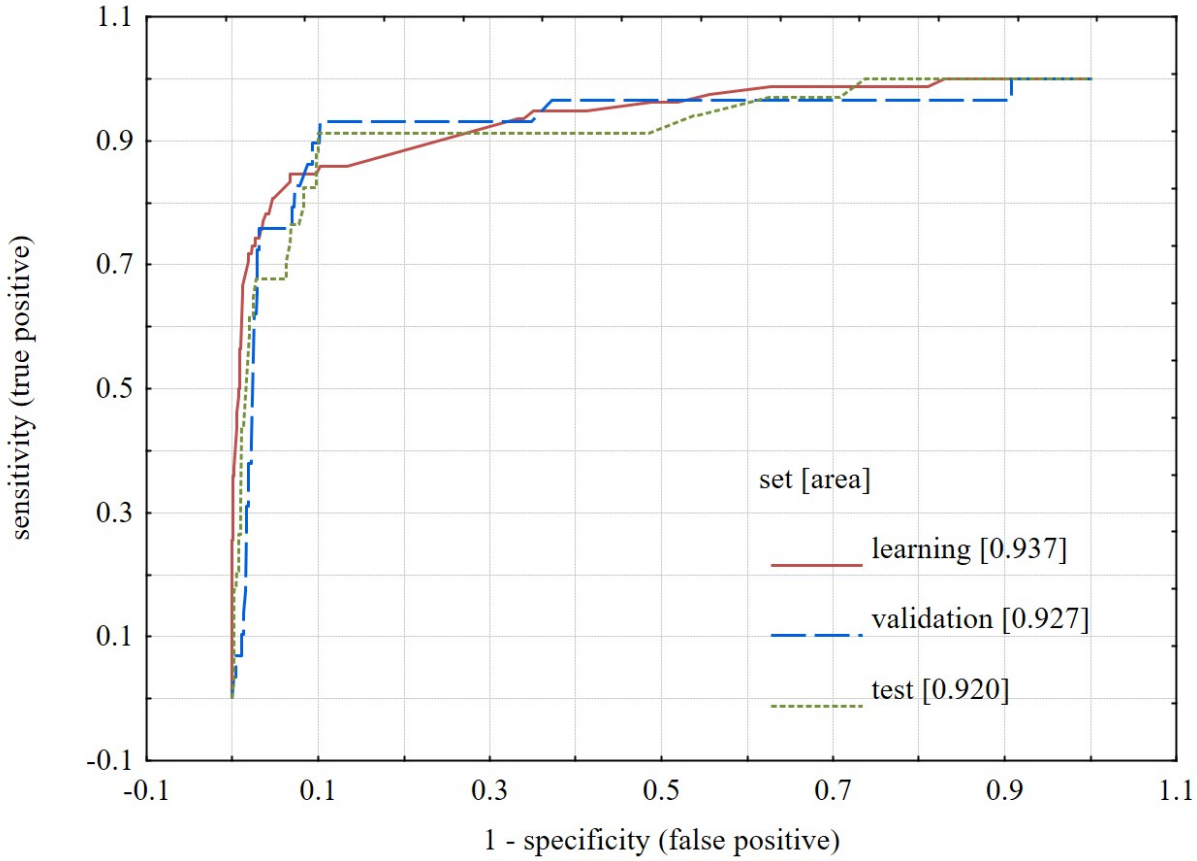

Figure 2. ROC curves for learning, test, and validation sets.

Comment 4. In Figure 2 Authors should distinguish the colors of the three plotted curves.

Answer:

The ROC curves for learning, validation and test sets on Figure 2 were distinguished with colors as reviewer suggested (above).

Comment 5. In lines 192-193 Authors say, "The selection of predictors was made based on clinical importance for recurrence occurrence …". Indeed, in a Deep and Machine Learning models, the feature selection process is essential to avoid the curse of dimensionality, eliminating the redundant features and keeping the most importance ones. I'd suggest the Authors to expand this concept of the feature importance and selection introducing the modern Explainable Artificial Intelligence (XAI) to overcome the black box nature of the Deep and Machine Learning models. In this sense, I'd invite the Authors to also include these references:

• N. Amoroso et al., "A roadmap towards breast cancer therapies supported by explainable artificial intelligence", Appl. Sci. 11, 4881 (2021).

• D. Gunning et al., "XAI—Explainable artificial intelligence", Sci. Robot. 4, 37 (2019).

Answer:

According to reviewer's suggestion, the concept of Explainable Artificial Intelligence (XAI) has been introduced in the section of variable selection and importance:

In order for AI to be effectively applied in the medical field, it is essential for clinicians to be able to interpret the data and outcomes. Therefore, it is important to carefully and thoughtfully select the input variables. This is where Explainable Artificial Intelligence (XAI) comes in, which places greater emphasis on human comprehension of AI processes rather than just strict calculations. This approach makes it easier to understand and apply the models to real-life problems.(20,21) (line 196-201)

References:

20. Amoroso N, Pomarico D, Fanizzi A, Didonna V, Giotta F, La Forgia D, et al. A roadmap towards breast cancer therapies supported by explainable artificial intelligence. Appl Sci. 2021 Jun 1;11(11):4881.

21. Gunning D, Stefik M, Choi J, Miller T, Stumpf S, Yang GZ. XAI-Explainable artificial intelligence. Sci Robot. 2019 Dec 18;4(eaay7120).

6. In the line 201 of section ANN analysis, various learning algorithms are mentioned (e.g., fast propagation, back propagation, quasi-Newton Broyden-Fletcher-Goldfarb-Shanno and conjugate gradients) but could the Authors, please, provide suitable references?

Answer:

Following the Reviewer's suggestion, the references were added.

The models learned with four learning algorithms – quickprop, backpropagation, quasi-Newton Broyden–Fletcher–Goldfarb–Shanno (BFGS), and conjugate gradients (22–24). (line 208-210)

References

22. Kelley CT, Society for Industrial and Applied Mathematics. 4. The BFGS Method, Iterative Methods for Optimization. In: Frontiers in Applied Mathematics [Internet]. 1999 [cited 2023 Jul 23]. p. 71–86. Available from: https://doi.org/10.1137/1.9781611970920.ch4

23. Nazareth JL. Conjugate gradient method. Wiley Interdiscip Rev Comput Stat. 2009 Nov;1(3):348–53.

24. Manning T, Sleator RD, Walsh P. Biologically inspired intelligent decision making: A commentary on the use of artificial neural networks in bioinformatics. Bioengineered. 2014 Dec 16;5(2):80–95.

Comment 7. In the Discussion section, Authors discuss their results. Could it be possible to create a Table to compare their results with the past literature of the topic?

Answer:

In accordance with the Reviewer's and referring to a similar comment of the Reviewer 1, part of the discussion was modified and a table comparing data from the literature from the last 5 years was attached (line 317 -339, Table 5)

The strategy of management and treatment of patients with lung cancer has undergone significant changes in recent years. Many years of clinical research, including preclinical and laboratory studies, have resulted in a modern approach to the oncological treatment. The optimal management strategy for NSCL is radical surgical resection in patients with stage I and II tumors (32). Patients not qualified or refusing this method of treatment should be treated with radical radiotherapy. Chemoradiation is the preferred treatment for patients with stage II and III disease who are not eligible for surgery. The optimal treatment strategy for stage N2-3A/3B non-small cell lung cancer (NSCLC) remains controversial due to its heterogeneity. Karacz et all, in a paper published in 2021, characterized 1549 cancer patients in terms of time of onset, location and recurrence in patients with lung cancer. The authors of this study showed that cases treated with systemic therapy had a higher rate of recurrence (47).

For patients with stage IV disease, there are still many uncertainties as to whether local consolidation therapy with radiotherapy or surgery for residual disease can improve overall survival. In the presented study, we collected data on patients hospitalized in the years 2012-2017, so that the 5-year observation period was maintained. Therefore, the reports presented in the literature look ahead to the future in relation to the patients presented by us. Nevertheless, based on the data available in the literature, it can be concluded that specific monitoring of patients and combination treatment of chemotherapy with immunotherapy brings significant benefits in some patients. In the Table 5 a brief review of the literature from the last 5 years on the management of patients with NSCLC is presented.

Table 5: Literature review

No.	Title	Stage	Group size	Conclusion	References Data
1.	Timing, sites, and correlates of lung cancer recurrence	I to III	1,549	cases treated with systemic therapy had a higher rate of recurrence	(47)
2.	Dynamic recurrence risk and adjuvant chemotherapy benefit prediction by ctDNA in resected NSCLC	II to III	116	Positive ctDNA results after both surgery and adjuvant chemotherapy are significantly associated with poorer recurrence-free survival.	(48)

3.	Neoadjuvant Nivolumab plus Chemotherapy in Resectable Lung Cancer	IB to IIIA	505	The addiction nivolumab to chemotherapy resulted in significantly longer event-free survival and a higher percentage of patients with a pathological complete response than chemotherapy alone.	(49)
4.	Postoperative intensity-modulated radiation therapy reduces local recurrence and improves overall survival in III-N2 non-small-cell lung cancer: A single-center, retrospective study	III-pN2	183	Postoperative Radiotherapy (PORT) significantly reduced the frequencies of local recurrence and improved overall survival (OS) in stage III-pN2 NSCLC, especially in the multiple-station pN2 group	(50)
5.	Osimertinib in Resected EGFR-Mutated Non-Small-Cell Lung Cancer	IB to IIIA	*682*	*EGFR* mutation-positive NSCLC, disease-free survival was significantly longer among those who received osimertinib	(51)
6.	Effect of Postoperative Radiotherapy for Patients With pIIIA-N2 Non-Small Cell Lung Cancer After Complete Resection and Adjuvant Chemotherapy: The Phase 3 PORT-C Randomized Clinical Trial	pIIIA-N2	394	Postoperative radiotherapy did not increase disease-free survival (DSF)	(52)
7.	Five Year Survival Update From KEYNOTE-010: Pembrolizumab Versus Docetaxel for Previously Treated, Programmed Death-Ligand 1-Positive Advanced NSCLC	IIIB/IV	1033	The use of pembrolizumab has a positive effect on overall survival (OS) and progression free survival (PFS) compared to docetaxel in patients with previously treated advanced NSCLC with PD-L1 expression.	(53)
8.	Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer	IIA to IV	133	Patients' monitoring during chemotherapy contributes to longer survival time and increasing the chances of detecting recurrence.	(54)
9.	Randomized phase II study of pemetrexed or pemetrexed plus bevacizumab for elderly patients with previously untreated non-squamous non-small cell lung cancer: Results of the Lung Oncology Group in Kyushu (LOGIK1201)	IIIB to IV	51	The addition of Bevacizumab to Pemetrexed does not improve survival in elderly patients with NSqNSCLC.	(55)
10.	Randomized phase II study of daily versus alternate-day administrations of S-1 for the elderly patients with completely resected pathological stage IA (tumor diameter > 2 cm)-IIIA of non-small cell lung cancer: Setouchi Lung Cancer Group Study 1201	IA to IIIA	101	Postoperative adjuvant chemotherapy for NSCLC is associated with improved survival in the elderly population.	(56)
11.	International Tailored Chemotherapy Adjuvant	II-IIIA	773	There was no statistical difference in OS in the use of adjuvant chemotherapy in	(57)

	(ITACA) trial, a phase III multicenter randomized trial comparing adjuvant pharmacogenomic-driven chemotherapy versus standard adjuvant chemotherapy in completely resected stage II-IIIA non-small-cell lung cancer			patients after complete resection of NSCLC.	
12.	Survival benefit of using pemetrexed for EGFR mutation-positive advanced non-small-cell lung cancer in a randomized phase III study comparing gefitinib to cisplatin plus docetaxel (WJTOG3405)		144	Sequential treatment including pemetrexed against EGFR-mutated NSCLC contributes to significantly longer OS	(58)
13.	Final overall survival results of WJTOG3405, a randomized phase III trial comparing gefitinib versus cisplatin with docetaxel as the first-line treatment for patients with stage IIIB/IV or postoperative recurrent EGFR mutation-positive non-small-cell lung cancer	IIIB/IV	177	The OS advantage of patients treated with gefitinib in a first line treatment compared to patients treated with cisplatin plus docetaxel has not been demonstrated. OS of patients with postoperative recurrence was better than that of stage IIIB/IV disease,	(59)
14.	Randomized Phase II Study of 3 Months or 2 Years of Adjuvant Afatinib in Patients With Surgically Resected Stage I-III EGFR-Mutant Non-Small-Cell Lung Cancer	IA to IIIB	46	Long-term (2-year) treatment with afatinib contributes to a reduction in the recurrence compared to shorter (3-month) treatment.	(60)
15.	Machine learning application in personalised lung cancer recurrence and survivability prediction	IA to IV (with no subdivision stages)	998	Good predictiveness of CART models were demostrated.	(61)

Comment: 8. Always in the Discussion section or eventually in the Introduction, Authors could mention these papers about non-small cell lung cancer (NSCLC) prediction, as a form of lung cancer, to mention past studies about Deep and Machine Learning methods for the topic of lung cancer recurrence:

• S. Hindocha et al., "A comparison of machine learning methods for predicting recurrence and death after curative-intent radiotherapy for non-small cell lung cancer: Development and validation of multivariable clinical prediction models", EBioMedicine 77, 103911 (2022).
• P. Aonpong et al., "Genotype-Guided Radiomics Signatures for Recurrence Prediction of Non-Small Cell Lung Cancer", IEEE Access 9, 90244-90254 (2021).
• G. Kim et al., "Deep Learning with Multimodal Integration for Predicting Recurrence in Patients with Non-Small Cell Lung Cancer", Sensors 22, 6594 (2022).
• X. Wang et al. "Prognostic recurrence analysis method for non-small cell lung cancer based on

CT imaging", International Conference on Image and Video Processing, and Artificial Intelligence. SPIE (2019).

Answer:

According to Reviwers suggestions the Introduction and Discussion part were revised and supplemented with literature data.

References:

37. Wang X, Duan H hong, Nie S dong. Prognostic recurrence analysis method for non-small cell lung cancer based on CT imaging. Proc SPIE 11321, 2019 International Conference on Image and Video Processing, and Artificial Intelligence [Internet]. 2019 Nov 27 [cited 2023 Jul 19];113211T:411–7. Available from: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11321/113211T/Prognostic-recurrence-analysis-method-for-non-small-cell-lung-cancer/10.1117/12.2539428.full

38. Aonpong P, Iwamoto Y, Han XH, Lin L, Chen YW. Genotype-Guided Radiomics Signatures for Recurrence Prediction of Non-Small Cell Lung Cancer. IEEE Access. 2021;9:90244–54.

39. Yang Y, Xu L, Sun L, Zhang P, Farid SS. Machine learning application in personalised lung cancer recurrence and survivability prediction. Comput Struct Biotechnol J. 2022 Jan 1;20:1811–20.

40. Hindocha S, Charlton TG, Linton-Reid K, Hunter B, Chan C, Ahmed M, et al. A comparison of machine learning methods for predicting recurrence and death after curative-intent radiotherapy for non-small cell lung cancer: Development and validation of multivariable clinical prediction models. EBioMedicine [Internet]. 2022 Mar [cited 2023 Jul 23];77(103911). Available from: https://doi.org/10.1016/j.ebiom.2022.103911

41. Kim G, Moon S, Choi JH. Deep Learning with Multimodal Integration for Predicting Recurrence in Patients with Non-Small Cell Lung Cancer. Sensors [Internet]. 2022 Aug 31 [cited 2023 Jul 19];22(17):6594. Available from: https://www.mdpi.com/1424-8220/22/17/6594/htm

Comment: The purpose of this study was to construct an ANN model based on medical records to predict lung cancer recurrence in patients with histologically diagnosed lung cancer. The authors report high prediction performance of 89.1% and 89.9% in the validation and test sets.

First, the data are not well presented. In addition to the fact that all lung cancer patients are included, the c-Stage is not presented. Since the objective is to predict recurrence, the initial treatment should be investigated in detail. However, this information is also lacking in this study. For example, whether complete resection was achieved by surgery or CR was achieved by chemotherapy. Since the data are incomplete or even if complete, they are not adequately presented in the first place, I think that the analytical results obtained from them are not sufficiently reliable.

Answer:

Thank you for the feedback on this study. You pointed out significant limitations of our work. Our aim was to find if the ANNs are able to predict the recurrence at all, even in that heterogenous group of patients, which in our opinion was achieved. The data were complete, but we are aware that we need to focus on more detailed clinical data, to achieve better model performance. We treat this research as preliminary research that shows the capability of the Neural Networks in the field of lung oncology.

Comment: For unbalanced data, performance cannot be evaluated by accuracy. Simply predicting all cases as having recurrence yields an accuracy of 93.2%, 94.9%, and 94.1% for the respective sets. Of course, the AUC has also been shown to be high, which is not to say that it is worthless, but if it were fully understood, it would not be presented in this way.

Answer:

Thank you for this valuable comment. Of course, the dataset is highly unbalanced and the classification cannot be described simply as ACC. As the Reviewer suggested, the classification metrics were revised, suited to the dataset (Table 3) , and included in the methods and discussion section (line 264-280)

Comment: And since the analysis is not categorized by treatment, it is unclear what benefit can be derived from the results. For example, the postoperative recurrence rate of lung cancer varies widely by stage, but postoperative follow-up is often uniform. It would be useful for surgeons if the recurrence rate of lung cancer could be determined more accurately than stage. The same is true for medical treatment.

Answer:

This analysis has been conducted as preliminary, and it will be further developed in the future. In the presented study presented, we took into account the data of 2,296 patients hospitalized in our center. Only 39 patients underwent surgical treatment. It constitutes only 1.70%. According to the presented sensitivity analysis, surgical treatment was ranked by 8. The vast majority of our patients were treated with chemotherapy (rank 1), radiotherapy (rank 2) or received molecular treatment (rank 3). In connection with the above, in our opinion, such a small percentage of pateints treated surgically does not affect the overall result. Nevertheless, we fully agree with the reviewer that a separate evaluation of surgical patients would be advisable. Our study, however, does not provide such an opportunity. We added this information about in the limitation section. The suggestions provided by the reviewer have been greatly appreciated and will be integrated into our future work.

Comment: In the discussion, where surgery is discussed, it is stated that "in the past, it was thought that the smaller the resection, the better for the patient." This is different from actual lung cancer practice.

Answer:

The discussion was corrected according to the reviewers' comments, a table comparing our results with the data available in the literature was added (Table 5)

Comment: The analysis method is also questionable: MLP is a classical neural network. There are several machine learning methods for table data, but in general, random forest and gradient boost tend to perform well. Since this study is aimed at prediction, there seems to be no reason not to use the above models, which have excellent prediction performance.

Answer:

In this article, the authors decided to use the classical MLP neural network as one of the machine learning methods that can be used to predict the outcome but we are aware that other predictive methods can be applied to solve the problem. The authors appreciate any recommendations and plan to expand their research to incorporate these methods in the future.