Accepted manuscript at https://doi.org/10.1007/978-3-031-49068-2_37

Assessing the performance of MRI-radiomic prognostic signatures in head and neck cancer patients: a comparative analysis

Anna Corti[1[0000-0001-9603-7825]], Luca Mainardi[1[0000-0002-6276-6314]], Valentina Corino[1,2[0000-0003-1825-0422]]

[1] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy
[2] Cardiotech Lab, Centro Cardiologico Monzino IRCCS, Milan, Italy
anna.corti@polimi.it

Abstract. The number of radiomic studies has dramatically increased in the last decade. However, the reproducibility of radiomic analyses remains challenging, although being fundamental for their clinical transferability. Herein we proposed an analysis of published radiomic signatures based on magnetic resonance imaging (MRI) for prognosis of overall survival (OS) in head and neck cancer (HNC) patients. Specifically, 5 reproducible MRI-based radiomic signatures were identified and their performance was tested on an external dataset of n=137 HNC patients. Although the lack of complete methodological details was encountered in the analyzed radiomic studies, the analysis performed herein allowed identifying 3 over 5 significantly prognostic signatures for OS of HNC patients. In particular, the radiomic signature with the highest stratification capability between high/low risk groups, provided a C-index 0.60, HR 2.43 and log-rank $p = 0.0022$. Overall, the study demonstrated the feasibility of comparing published radiomic signatures on an external dataset (provided that sufficient methodological details are reported). In future, major efforts should be put in reporting radiomic analyses in order to enable their full reproduction in view of their potential translation in clinics.

Keywords: Radiomics, Overall survival, Head and neck cancer.

1 Introduction

Head and neck squamous cell carcinoma (HNC) are a group of highly heterogeneous malignancies representing the seventh most common and the sixth most deadly tumor worldwide (accounting for more than 350000 annual deaths) [1]. Nowadays, the tumor-node-metastasis (TNM) staging system is the main factor guiding risk assessment, treatment choice and prognosis [2]. However, the stratification performance of staging-based system is quite low. The high heterogeneity of HNC and the emergence of personalized medicine fostered the development of additional biomarkers. Radiomics, namely the quantitative extraction of high throughput data from medical images, demonstrated to hold the potential for providing HNC biomarkers [3]. The number of

Accepted manuscript at https://doi.org/10.1007/978-3-031-49068-2_37

radiomic studies has dramatically increased in the last decade. In this context, the repeatability and reproducibility of radiomic analyses remains an open issue, which needs to be addressed in order to enable the clinical translation of radiomics [4].

The aim of the present study was to test the prognostic performance of reproducible radiomic signatures for prognosis of overall survival (OS) based on magnetic resonance imaging (MRI) on a HNC patient dataset.

2 Methods

2.1 Patient dataset, MRI image preprocessing and features extraction

A subset of locally advanced HNC (clinical TNM III/IV according to the 8th edition of AJCC/UICC) of the BD2Decide project (NCT02832102) presenting with pre-contrast T1-weighted (T1w) and T2-weighted (T2w) and post-contrast T1w (T1wCont) MR image sequence (acquired with 1.5 T scanner) was used in the present study (n=137 patients) [5].

The region of interest (ROI) corresponding to the primary tumor volume was segmented by an expert radiologist. MRI images underwent the following preprocessing steps: (i) denoising, through a 3D Gaussian filter with a 3x3x3 voxel kernel and σ = 0.5; (ii) intensity non-uniformities correction, through the N4ITK algorithm [6]; (iii) intensity standardization, through Z-score; (iv) voxel size resampling, through B-spline interpolation. Radiomic features were extracted from the original image and transformed images, including the Laplacian of Gaussian (σ=0.5, 1.0, 1.5, 2.0 and 5.0 mm) the wavelet, the square, the square root and the logarithm filters. For each original and transformed image, features belonging to first order statistics, shape and size (only for original images), grey level cooccurrence matrix, grey level size zone matrix, neighboring gray tone difference matrix and grey level dependence matrix were extracted, for a total of n = 5064 features. Pyradiomics 2.2.0 software was used to extract the features [7].

2.2 Radiomic signature testing

A literature survey was performed to retrieve reproducible MRI-radiomic signatures for prognosis of OS in HNC patients and compute the radiomic scores on our HNC dataset. To reproduce the signatures, the methods applied in the studies were followed. However, if information regarding features normalization and/or signature dichotomization was missing, the following criteria were applied. As regards features normalization, (i) if the information was missing, the original features were considered, (ii) if features normalization was mentioned, but without providing additional details (e.g., method and features median/standard deviation), Z-score normalization was applied on our data. Similarly, to dichotomize the radiomic signatures (between high/low risk groups), if the threshold was not provided, the median value of the radiomic signature in the present data was used.

The identified reproducible radiomic signatures were thus computed on our dataset. The prognostic performance of the radiomic signatures was compared in terms of

Accepted manuscript at https://doi.org/10.1007/978-3-031-49068-2_37

Harrel's concordance index (C-index), hazard ratio (HR) and high/low risk group patient stratification with Kaplan-Meier curves.

3 Results

3.1 Survey of literature

Five reproducible MRI-based radiomic studies [8–12], providing information on the selected features along with regression coefficients, were identified and reported in Table 1. Four monomodal and one multimodal radiomic signatures were reported. Most of the radiomic signatures were based on T1wCont sequence (Rad-1, Rad-2, and Rad-4), one radiomic signature was based on T2w sequence (Rad-5) and one multimodal radiomic signature was based on the T1w, T1wCont and T2w sequences (Rad-3). The number of the selected features span from 2 to 10. Moreover, the selected features were all different among the studies, leading to a total of 25 potentially prognostic features identified (i.e., 21 from T1wCont, 3 from T2w and 1 from T1w). Specifically, (i) among the 21 T1wCont features, 13 were extracted from the wavelet transformation (textural features), 5 from the Laplacian of Gaussian transformation (3 first order and 2 textural features) and 3 from the original image (2 shape and 1 textural features); (ii) among the 3 T2w features, one was extracted from the wavelet transformation (textural feature), one from the Laplacian of Gaussian transformation (first order feature) and one from the original image (shape feature) and (iii) the T1w feature was extracted from the original image (textural feature).

As regards features normalization, (i) to compute Rad-1, Z-score standardization was applied on the features, as reported in the study [8], (ii) to compute Rad-2 and Rad-4, the original features were considered because standardization processes were not mentioned in the corresponding studies [9, 11], (iii) to compute Rad-3, Z-score standardization was applied on the features, by considering the mean and standard deviation reported in the study [10], and (iv) to compute Rad-5, Z-score standardization was applied on the features, because feature normalization was mentioned in the study, but details on the normalization method were not provided [12].

Finally, as regards the cutoff value to dichotomize the signature, except for Rad-4 for which the cutoff value was available, the median value was considered.

Table 1. MRI-based radiomic signatures included in the study.

Signature	Study	Features	Feature normalization	Cutoff
Rad-1	Bos 2021 [8]	10 from T1wCont	Z-score – no details	NA
Rad-2	Chen 2022 [9]	6 from T1wCont	NA	NA
Rad-3	Alfieri 2022 [10]	3 from T1w, T1wCont, T2w	Z-score – details	NA
Rad-4	Siow 2022 [11]	4 from T1wCont	NA	0.5
Rad-5	Mossinelli 2023 [12]	2 from T2w	Standardized – no details	NA

Accepted manuscript at https://doi.org/10.1007/978-3-031-49068-2_37

3.2 Radiomic features testing

Table 2 lists the results of the radiomic signatures and the combined radiomic score and Fig.1 shows the corresponding Kaplan-Maier curves. Three over five radiomic signatures (i.e., Rad-2, Rad-3 and Rad-5) were significantly prognostic for the OS of the considered dataset, with Rad-5 providing the best performance, with C-index 0.60 (IQR 0.57-0.63), HR 2.43 (95% confidence interval 1.41-4.20) and log-rank p = 0.0022.

Table 2. Performance of radiomic signatures.

Signature	C-index	Log-rank HR	Log-rank *p*
Rad-1	0.56 [IQR 0.53 0.59]	1.40 [95% CI 0.82 2.41]	0.28
Rad-2	0.55 [IQR 0.52 0.57]	1.78 [95% CI 1.04 3.06]	0.05
Rad-3	0.60 [IQR 0.58 0.62]	2.06 [95% CI 1.20 3.55]	0.013
Rad-4	0.48 [IQR 0.45 0.51]	0.92 [95% CI 0.24 3.61]	0.82
Rad-5	0.60 [IQR 0.57 0.63]	2.43 [95% CI 1.41 4.20]	0.0022

C-index: Harrel's concordance index; HR: hazard ratio

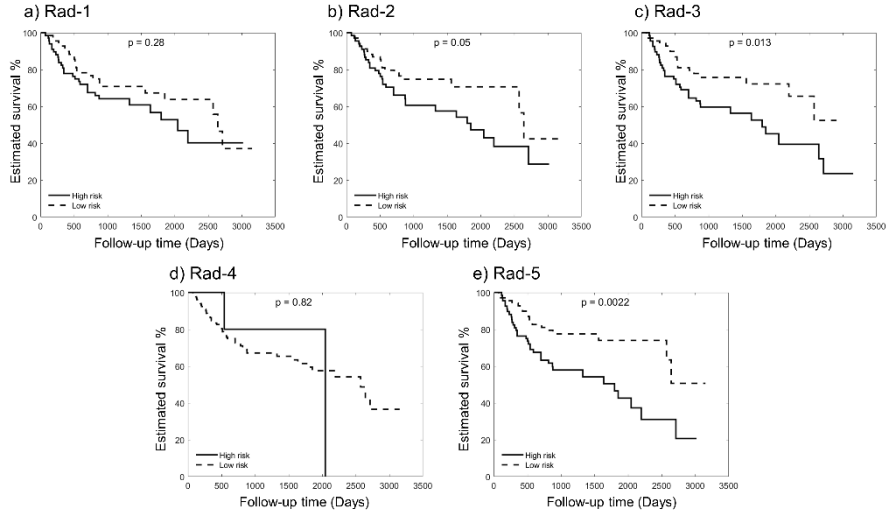

Fig. 1. Kaplan-Meier curves for (a) Rad-1, (b) Rad-2, (c) Rad-3, (d) Rad-4 and (e) Rad-5.

4 Discussion

The number of radiomic studies has dramatically increased in the last decade. However, the lack of common consensus in the applied methodologies and the paucity of details reported in the studies, make the reproducibility of published results a challenge. This in turn hampers the potential translation of radiomics to clinics [4]. The present study proposed an analysis of published MRI-based radiomic signatures for prognosis of OS

in HNC patients, with the goal of testing their replicability and performance on an external dataset of n=137 HNC patients.

Overall, a faithful reproduction of the signatures was not possible because of the lack of complete methodological details (e.g., either about feature normalization or cutoff value for dichotomization). This sheds light on the need for defining a common consensus about the transparency of the delivered information, which is required to correctly replicate the radiomic analyses and subsequently to lay the foundations for a potential clinical translation. However, to overcome the lack of details and compute the radiomic signatures, some assumptions were introduced. With the exception of Rad-4, all the radiomic signatures separated the high- from the low-risk groups, with 3 of them presenting with significant p-value of the log-rank test.

The present study is not exempt from limitations, mainly associated to the necessary assumptions that were considered in order to compute the radiomic signatures. In future, performing a meta-analysis of HNC radiomic prognostic signatures is thought to provide a remarkable impact in the field.

5 Conclusion

This study demonstrated the feasibility of comparing published radiomic signatures on an external dataset, provided that sufficient methodological details are described. Future efforts should be put in reporting radiomic analyses in order to enable their full reproduction in view of their potential translation in clinics.

References

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A.: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. 68, 394–424 (2018). https://doi.org/10.3322/caac.21492.
2. Machiels, J.-P., René Leemans, C., Golusinski, W., Grau, C., Licitra, L., Gregoire, V.: Squamous cell carcinoma of the oral cavity, larynx, oropharynx and hypopharynx: EHNS-ESMO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up., (2020). https://doi.org/10.1016/j.annonc.2020.07.011.
3. Bruixola, G., Remacha, E., Jiménez-Pastor, A., Dualde, D., Viala, A., Montón, J.V., Ibarrola-Villava, M., Alberich-Bayarri, Á., Cervantes, A.: Radiomics and radiogenomics in head and neck squamous cell carcinoma: Potential contribution to patient management and challenges. Cancer Treat. Rev. 99, 102263 (2021). https://doi.org/10.1016/j.ctrv.2021.102263.
4. Pfaehler, E., Zhovannik, I., Wei, L., Boellaard, R., Dekker, A., Monshouwer, R., El Naqa, I., Bussink, J., Gillies, R., Wee, L., Traverso, A.: A systematic review and quality of reporting checklist for repeatability and reproducibility of radiomic features. Phys. imaging Radiat. Oncol. 20, 69–75 (2021). https://doi.org/10.1016/j.phro.2021.10.007.
5. Cavalieri, S., De Cecco, L., Brakenhoff, R.H., Serafini, M.S., Canevari, S., Rossi, S., Lanfranco, D., Hoebers, F.J.P., Wesseling, F.W.R., Keek, S., Scheckenbach, K., Mattavelli, D., Hoffmann, T., López Pérez, L., Fico, G., Bologna, M., Nauta, I., Leemans, C.R., Trama,

A., Klausch, T., Berkhof, J.H., Tountopoulos, V., Shefi, R., Mainardi, L., Mercalli, F., Poli, T., Licitra, L.: Development of a multiomics database for personalized prognostic forecasting in head and neck cancer: The Big Data to Decide EU Project. Head Neck. 1–12 (2020). https://doi.org/10.1002/hed.26515.

6. Tustison, N.J., Cook, P.A., Gee, J.C.: N4ITK: improved N3 bias correction. IEEE Trans. Med. Imaging. 29, 1310–1320 (2010). https://doi.org/10.1109/TMI.2010.2046908.N4ITK.
7. van Griethuysen, J.J.M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G.H., Fillion-Robin, J.-C., Pieper, S., Aerts. Hugo J. W. L.: Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res. 77, e104–e107 (2017).
8. Bos, P., van den Brekel, M.W.M., Gouw, Z.A.R., Al-Mamgani, A., Taghavi, M., Waktola, S., Aerts, H.J.W.L., Castelijns, J.A., Beets-Tan, R.G.H., Jasperse, B.: Improved outcome prediction of oropharyngeal cancer by combining clinical and MRI features in machine learning models. Eur. J. Radiol. 139, 109701 (2021). https://doi.org/10.1016/j.ejrad.2021.109701.
9. Chen, J., Lu, S., Mao, Y., Tan, L., Li, G., Gao, Y., Tan, P., Huang, D., Zhang, X., Qiu, Y., Liu, Y.: An MRI-based radiomics-clinical nomogram for the overall survival prediction in patients with hypopharyngeal squamous cell carcinoma: a multi-cohort study. Eur. Radiol. 32, 1548–1557 (2022). https://doi.org/10.1007/s00330-021-08292-z.
10. Alfieri, S., Romanò, R., Bologna, M., Calareso, G., Corino, V., Mirabile, A., Ferri, A., Bellanti, L., Poli, T., Marcantoni, A., Grosso, E., Tarsitano, A., Battaglia, S., Blengio, F., De Martino, I., Valerini, S., Vecchio, S., Richetti, A., Deantonio, L., Martucci, F., Grammatica, A., Ravanelli, M., Ibrahim, T., Caruso, D., Locati, L.D., Orlandi, E., Bossi, P., Mainardi, L., Licitra, L.F.: Prognostic role of pre-treatment magnetic resonance imaging (MRI)-based radiomic analysis in effectively cured head and neck squamous cell carcinoma (HNSCC) patients. Acta Oncol. 60, 1192–1200 (2021). https://doi.org/10.1080/0284186X.2021.1924401.
11. Siow, T.Y., Yeh, C.-H., Lin, G., Lin, C.-Y., Wang, H.-M., Liao, C.-T., Toh, C.-H., Chan, S.-C., Lin, C.-P., Ng, S.-H.: MRI Radiomics for Predicting Survival in Patients with Locally Advanced Hypopharyngeal Cancer Treated with Concurrent Chemoradiotherapy. Cancers (Basel). 14, (2022). https://doi.org/10.3390/cancers14246119.
12. Mossinelli, C., Tagliabue, M., Ruju, F., Cammarata, G., Volpe, S., Raimondi, S., Zaffaroni, M., Isaksson, J.L., Garibaldi, C., Cremonesi, M., Corso, F., Gaeta, A., Emili, I., Zorzi, S., Alterio, D., Marvaso, G., Pepa, M., De Fiori, E., Maffini, F., Preda, L., Benazzo, M., Jereczek-Fossa, B.A., Ansarin, M.: The role of radiomics in tongue cancer: A new tool for prognosis prediction. Head Neck. (2023). https://doi.org/10.1002/hed.27299.