Head and Neck Cancer Overall Survival Prognostication Using Dosiomic Features and Random Survival Forest Algorithm

Zahra Mansouri[1], Yazdan Salimi[1], Mehdi Amini[1], Ghasem Hajianfar[2], Mehrdad Oveisi[3], Isaac Shiri[1], Habib Zaidi, *FIEEE*[1,4,5,6]

[1]Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva, Switzerland
[2] ajaie Cardiovascular, Medical & Research Center, Iran University of Medical Science, Tehran, Iran
[3]Comprehensive Cancer Centre, School of Cancer & Pharmaceutical Sciences, Faculty of Life Sciences & Medicine, King's College London, London, United Kingdom
[4]Geneva University Neurocenter, University of Geneva, 1205 Geneva, Switzerland
[5]Department of Nuclear Medicine and Molecular Imaging, University of Groningen, Groningen, Netherlands
[6]Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark

INTRODUCTION

- the use of quantitative features extracted from medical images (Radiomics) proved to be promising in patient outcome modeling and can be a strong contribution to decision support towards personalized cancer treatment.
- For the patients who undergo radiotherapy (RT), there will be a 3D dose distribution based on their treatment plan, which contains lots of valuable information. Traditionally, a part of this information can be shown on dose-volume histograms (DVHs). However, the DVHs cannot show spatial information by their nature .
- Encoding the spatial and statistical information of RT dose maps the same way as radiomics (called Dosiomics) allows producing a valuable set of quantitative dose features.
- Head and neck (H&N) cancer patients show a poor prognosis, and their overall survival (OS) can be affected by many factors . Some studies attempted to model OS using different machine learning (ML) algorithms. Random survival forest (RSF) has been shown to be more straightforward and reliable.
- This study aimed to investigate the radiomics and Dosiomics features value on the modeling of overall survival (OS) of head and neck (H&N) cancer patients by RSF machine learning algorithm

METHODS

Dataset

We included 240 H&N cancer patients from five different institutions whose data were collected on the TCIA database. After registration of CT images on their corresponding dose distribution maps, GTV segments were extracted.

Radiomics Workflow

Features were extracted using a MATLAB®-based package known as Standardized Environment for Radiomics Analysis (SERA). Overall , 215 features, including 79 first-order features and 136 three-dimensional texture features, were extracted. Due to the significant role of clinical features in the prognosis of H&N cancer patients, we used 13 clinical features (gender, age, histology, smoking status, HPV status, T-staging, N-staging, TNM staging, primary tumor site, treatment modalities) alongside the radiomic and Dosiomic features.

Prognostic Modeling workflow

five feature selection (FS) approaches were used to find relevant features: C-Index, Variable hunting (VH), Variable hunting Variable Importance (VH. VIMP), Minimal Depth (MD), and Mutual Information (MI). In our prognostic modeling framework, these FS methods, combined with the random survival forest (RSF) machine learning (ML) algorithm, were applied to the radiomics, Dosiomics, and dose/ CT feature sets. The hyperparameters were optimized by grid search. Then concordance indices (C-Indices) were reported as the model performance quantification measured by 5-fold cross-validation for OS prediction. MLR package in R 3.6.2 was used for hyper-parameter optimization, model training, and evaluation.

RESULTS

- ✓ An example of CT, and Dose images are illustrated in Figure 1.
- ✓ Table 1 demonstrates the all C_indices± SD for all the strategies and models.
- ✓ Figure 2. shows the higher C-indices distribution for the Dose strategy (Dosiomics) compared to radiomics and integration of CT-Dose strategies.
- ✓ the achieved C-Indices are all above 0.61 for all the strategies, however, It is well depicted that the Dose strategy showed the highest values of C-index from the combination of MD-RSF (C-index= 0.71), Then the combination of VH, VIMP –RSF (C-index= 0.70) for the radiomics strategy.
- ✓ In another strategy, we used a mixture of Dose and CT features, but the results did not show any superiority to the radiomics and Dosiomics strategies alone.
- ✓ Figure 3 depicts the Kaplan–Meier curve for the best model (combination of MD-RSF) in which shows a good separation between the high and low risk groups.

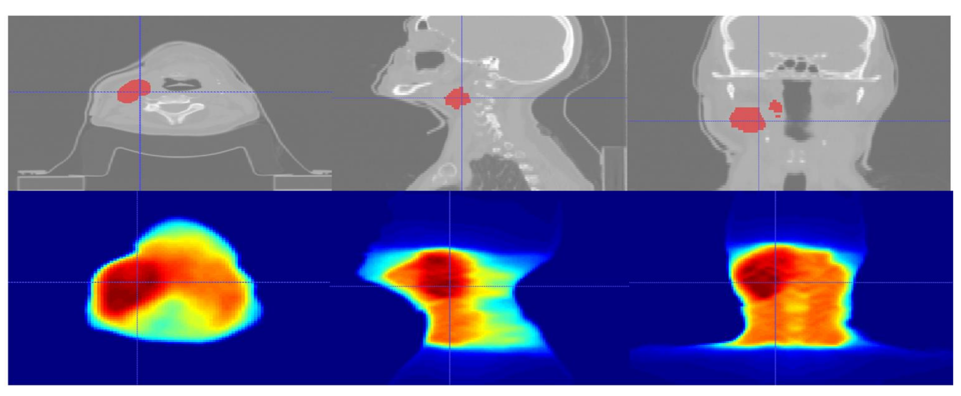

Figure 1. CT images with GTV segments (top panel) and dose maps (bottom).

Table1. C-indices of different models and strategies ± SDs

Strategy	CT	DOSE	Dual CT_Dose
RSF_Cindex	0.68 ± 0.11	0.67 ± 0.11	0.65 ± 0.095
RSF_MD	0.63 ± 0.092	0.71 ± 0.11	0.63 ± 0.091
RSF_MI	0.65 ± 0.1	0.66 ± 0.12	0.62 ± 0.091
RSF_VH.VIMP	0.7 ± 0.11	0.66 ± 0.12	0.63 ± 0.087
RSF_VH	0.66 ± 0.12	0.66 ± 0.11	0.61 ± 0.081

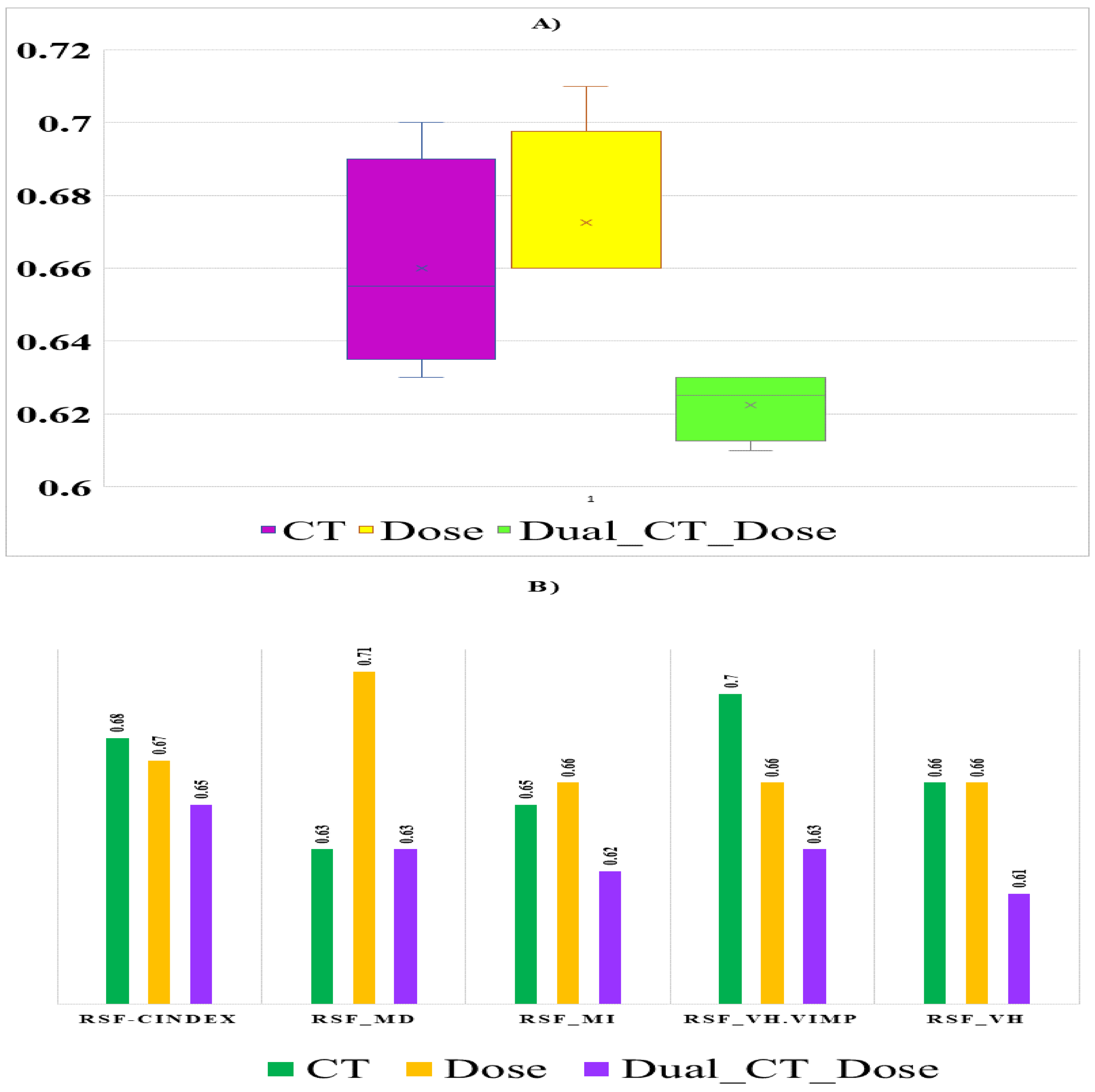

Figure 2. Plots showing the C-Indexes for the RSF machine learning algorithm in combination with different feature selection methods applied to CT (radiomics), Dose (Dosiomics) and, dual CT-Dose sets

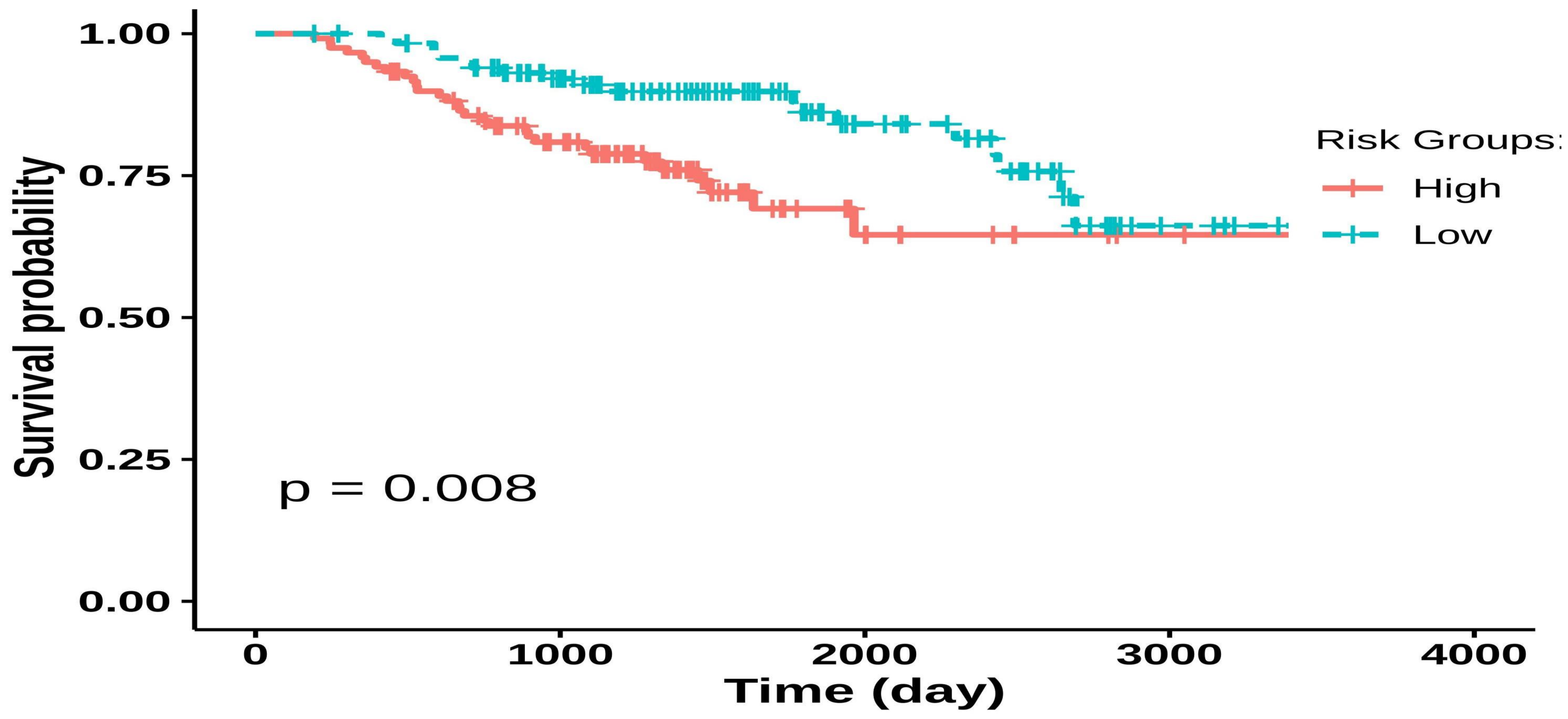

Figure 3. Kaplan–Meier curve (K-M curve) for the combination of MD-RSF model.

CONCLUSION

In this study, we used the RSF machine learning algorithm combined with five different FS methods to explore the value of Dosiomics and radiomics features in predicting the OS of H&N cancer patients after radiotherapy. Our findings showed that the radiotherapy 3D dose distribution contained important information highly correlated with the survival prediction of H&N cancer patients.

In conclusion, this study showed that RSF algorithms trained by Dosiomic and radiomic features could significantly predict the survival and prognostication of H&N cancer patients. At the same time, the Dosiomics strategy has a stronger role in predicting overall survival.

This work was supported by the Euratom research and training programme 2019-2020 Sinfonia project under grant agreement No 945196.