

PET-based prognostic survival model after radiotherapy for head and neck cancer

Joël Castelli, A Depeursinge, A Devillers, B Campillo-Gimenez, Y Dicente, J O Prior, E Chajon, F Jegoux, C Sire, O Acosta, et al.

▶ **To cite this version:**

Joël Castelli, A Depeursinge, A Devillers, B Campillo-Gimenez, Y Dicente, et al.. PET-based prognostic survival model after radiotherapy for head and neck cancer. European Journal of Nuclear Medicine and Molecular Imaging, 2019, 46 (3), pp.638-649. 10.1007/s00259-018-4134-9 . hal-01862535

HAL Id: hal-01862535

https://univ-rennes.hal.science/hal-01862535

Submitted on 30 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PET-based prognostic survival model after radiotherapy for head and neck cancer

Joël Castelli[1,2,3,*] & A. Depeursinge[4,5] & A. Devillers[6] & B. Campillo-Gimenez[2,3,7] & Y. Dicente[4,8] & J. O. Prior[9] & E. Chajon[1] & F. Jegoux[10] & C. Sire[11] & O. Acosta[2,3] & E. Gherga[12] & X. Sun[12,13] & B. De Bari[13] & J. Bourhis[14] & R. de Crevoisier[1,2,3]

* Joël Castelli j.castelli@rennes.unicancer.fr

Abstract

Purpose The aims of this multicentre retrospective study of locally advanced head and neck cancer (LAHNC) treated with definitive radiotherapy were to (1) identify positron emission tomography (PET)-^{18}F-fluorodeoxyglucose (^{18}F-FDG) parameters correlated with overall survival (OS) in a training cohort, (2) compute a prognostic model, and (3) externally validate this model in an independent cohort.

Materials and methods A total of 237 consecutive LAHNC patients divided into training (n = 127) and validation cohorts (n = 110) were retrospectively analysed. The following PET parameters were analysed: SUV_{Max}, metabolic tumour volume (MTV), total lesion glycolysis (TLG), and SUV_{Mean} for the primary tumour and lymph nodes using a relative SUV_{Max} threshold or an absolute SUV threshold. Cox analyses were performed on OS in the training cohort. The c-index was used to identify the highly prognostic parameters. A prognostic model was subsequently identified, and a nomogram was generated. The model was externally tested in the validation cohort.

Results In univariate analysis, the significant PET parameters for the primary tumour included MTV (relative thresholds from 6 to 83% and absolute thresholds from 1.5 to 6.5) and TLG (relative thresholds from 1 to 82% and absolute thresholds from 0.5 to 4.5). For the lymph nodes, the significant parameters included MTV and TLG regardless of the threshold value. In multivariate analysis, tumour site, p16 status, MTV35% of the primary tumour, and MTV44% of the lymph nodes were independent predictors of OS. Based on these four parameters, a prognostic model was identified with a c-index of 0.72. The corresponding nomogram was generated. This prognostic model was externally validated, achieving a c-index of 0.66.

Conclusions A prognostic model of OS based on primary tumour and lymph node MTV, tumour site, and p16 status was proposed and validated. The corresponding nomogram may be used to tailor individualized treatment.

Keywords Head and neck cancer · Nomogram · Prognostic score · PET · Radiotherapy

Introduction

The American Joint Committee on Cancer (AJCC) staging system based on primary tumour extension and nodal spread is generally used to estimate the prognosis and guide the therapy of head and neck cancer [1]. Chemoradiotherapy (CRT) is a standard treatment for non-resected or unresectable locally advanced head and neck cancer (LAHNC) [2–4]. A potential alternative standard treatment, especially when concomitant chemotherapy cannot be used, is to combine radiotherapy with cetuximab [5]. Despite these treatments, the prognosis remains relatively poor, and loco-regional recurrence can occur in up to 40% of patients, mostly occurring in the first 2 years after treatment [6].

The use of tumour volume based on morphological or functional imaging instead of strict anatomic extent may better identify patients with a worse prognosis. Indeed, ^{18}F- fluorodeoxyglucose (^{18}F-FDG) positron emission tomography/computed tomography (PET/CT) parameters, in particular metabolic tumour volume (MTV) and total lesion glycolysis (TLG), are correlated with clinical outcome in head and neck cancer [7–10].

Accepted Manuscript

However, few studies have compared different thresholds of MTV and/or TLG [11–16], and a large majority of studies have used an SUV_{Max} threshold of 40% [17] or a fixed SUV threshold of > 2.5 mg/ml [18]. Another limit is the lack of data regarding the impact of the tumour site on the prognostic value of PET, with some studies including only a specific tumour site, and others including various tumour sites without subgroup analysis. Moreover, these studies did not perform external validation of PET parameters to predict overall survival (OS). Finally, in addition, there is a clear need to generate a validated model to predict individual outcome for a given external patient.

In this context, with regard to oropharyngeal, hypopharyngeal, and oral cavity cancers treated with CRT or radiotherapy (RT)-cetuximab, our study aimed (1) to identify clinical and PET parameters predicting OS from an initial cohort of patients, (2) to generate a prognostic model of OS, and (3) to validate this prognostic model with a second independent cohort of patients.

Material and methods

Inclusion criteria

All consecutive patients treated with definitive concurrent CRT or RT and cetuximab for LAHNC between January 2010 and May 2017 from four different hospitals were retrospectively reviewed. Inclusion criteria were an age between 18 and 75 years, stage III or IV (AJCC 7th edition), no surgery before RT, no history of cancer, PET performed less than 8 weeks prior to RT, no metastasis at diagnosis, and a minimal follow-up of 3 months. Nasopharyngeal cancers were excluded from the study.

Patient characteristics and treatment results

This retrospective study included 237 patients. The main patient, tumour, and treatment characteristics are presented in Table 1. All tumours were locally advanced, corresponding to T3-4 or N2-3 stage (stage III, IVa or IVb, AJCC 7th edition). Human-papilloma-virus status was determined using p16 status evaluated by immunohistochemistry. P16 status was available for 60% of the patients.

All patients underwent intensity-modulated RT (IMRT) using volumetric modulated arc therapy (VMAT) (73%) or helical tomotherapy (27%). Planning CTs with intravenous contrast agents were acquired with 2-mm slice thickness from the vertex to the carina. A thermoplastic head and shoulder mask with five fixation points was used. PET/CT and MRI co- registration was used for tumour delineation. A total dose of 70 Gy [2 Gy/fraction/day, 35 fractions (73%) or 2.12 Gy/fractions/day, 33 fractions (27%)] with a simultaneous integrated boost technique [19] was administered in combination with concomitant chemotherapy (platinum, 100 mg/m^2 thrice weekly [3, 4] or cetuximab at an initial dose of 400 mg/m^2 followed by 250 mg/m^2 weekly for the duration of radiotherapy [5] if the patient was not fit for chemotherapy) (Table 1). Three target volumes were generated on the planning CT: CTV70, CTV63, and CTV56. The 70 Gy clinical target volume (CTV70) was equal to the gross tumour volume plus a 5- mm 3D margin adjusted to exclude the air cavities and all bone mass free of tumour invasion. CTV63 (optional) corresponded to the area at high risk of microscopic spread, while CTV56 corresponded to the low-risk subclinical area. The planning target volume (PTV) included the CTVs plus a 5-mm 3D margin limited to 3 mm from the skin surface to avoid the build-up region and therefore limit skin toxicity [20]. The minimum PTV covered by the 95% isodose line was 95%. The organs at risk were manually delineated accord- ing to the GORTEC group (the French group of radiation oncology for head and neck cancer). The dose constraints in the volume of interest were also set according to the GORTEC group. The study was approved by the institutional ethical committees (NCT02469922).

Physical evaluation and laryngoscopy were performed after RT every 3 months for the first 2 years and then every 6 months thereafter. A CT scan or a PET/CT was performed between 3 and 6 months after treatment. The database was locked on the 15th of December 2017.

The entire cohort was divided into a training cohort, including 127 patients from one hospital, and a validation cohort, including 110 patients from three different hospitals. The median follow-up for the training and the validation cohort were 45 months (ranging from 6 to 92 months) and 23 months (ranging from 3 to 57 months) respectively. For the entire population, the 2-year OS rate was 65% [95% confidence interval

Accepted Manuscript

(95% CI): 59–72%]. The follow-up and treatment results of the training and validation cohorts are displayed in Table 1.

PET/CT image acquisition and analysis

All patients underwent FDG PET/CT for staging at most 8 weeks before RT. For three cohorts, the patients fasted at least 4 h prior to injection of 4 Mbq/kg of ^{18}F-FDG (Flucis, CIS bio, Saclay, France). Blood glucose levels (limit < 150 mg/dl) were assessed prior to the injection of ^{18}F- FDG. If not contra-indicated, intravenous contrast agents were administered before CT scanning. After a 60-min uptake pe- riod of rest, patients were imaged using the Discovery ST PET/CT imaging system (General Electric Medical SystemsTM, Milwaukee, WI, USA) or the Siemens Biograph 6 True Point PET/CT scanner (Siemens Medical Solutions, Erlangen, Germany). First, CT (120 kV, 80 mA, 0.8 s rotation time, slice thickness 3.75 mm) was performed from the base of the skull to the mid-thigh. PET scanning was performed immediately after CT acquisition. Images were ac- quired from the base of the skull to the mid-thigh (3 min/bed position). PET images were reconstructed using an ordered-subset expectation maximization iterative reconstruction (OSEM) (two iterations, 28 subsets) and an iterative fully 3D (Discovery ST). CT data were used for attenuation correction. A similar protocol was used for the last cohort; however, a smaller injection of 3.5 Mbq/kg of ^{18}F-FDG was used with a slightly more recent system (Discovery D690 TOF PET/CT, General Electric Healthcare, Milwaukee, WI, USA) that allowed for shorter acquisition times (2 min/bed position). PET images acquired by this system were reconstructed after time-of-flight and point-spread-function recovery corrections.

For each patient, gross tumour volume-tumour (GTV-T) and nodal GTV (GTV-N) were manually segmented on each PET/CT by the same radiation oncologist. A region of interest (ROI) was computed by adding a 3D margin of 5 mm to GTV- T (ROI-T) and GTV-N (ROI-N).

A set of quantitative parameters based on SUV histograms were extracted from ROI-T and ROI-N in PET images using the QuantImage web service [21]. SUV_{Max} was first computed from ROI-T as the maximum SUV in the delineated volume. Various metabolic volumes were subsequently defined based on two segmentation methods: (i) an absolute threshold of SUV (ranging from 0 to 20, 0.5 by step) or (ii) a relative threshold of SUV_{Max} [0-100% (1% by step)]. Metabolic in- tensity parameters were computed using the two segmentation methods at each threshold for both ROI-T and ROI-N. Relative thresholds for ROI-N were computed based on SUV_{Max} of the primary tumour. SUV_{Mean} was computed for each threshold. MTV was computed as the metabolic volume of the segmented region in millilitres. TLG was computed as SUV_{Mean} × MTV of the corresponding thresholded region.

Statistical analysis

OS was calculated from the first day of RT to the date of death from any cause. Patients alive at the time of analysis were censored at the date of last follow-up. Loco-regional control (LRC) was calculated from the first day of RT to the date of first recurrence in the primary tumour and/or lymph node. Follow-up was calculated using a reverse Kaplan–Meier estimation [22]. OS and LRC estimations were computed using the Kaplan–Meier method, and a two-sided log-rank test was used to compare groups.

The analyses were performed as suggested in the TRIPOD statement [23]. In the first step, the analysis was performed in the training cohort only. A univariate Cox analysis was first performed on OS to assess the following parameters: gender, age, smoking status, T and N classification, tumour site, p16 status, chemotherapy, GTV (in ml), SUV_{Max}, SUV_{Mean}, MTV, and TLG. Interactions between the tumour site and PET parameters were also investigated. The significant parameters were identified, and the Harrell's c-index was calculated (c- index ≈ 0.5 → random predictions, c-index ≈ 1 → perfect pre- diction) [24]. Among the different thresholds used to compute PET parameters, the highly prognostic parameter based on the c-index was selected for the MTV, TLG and SUV_{Mean}. Factors with significance of p value < 0.1 and with highest c-index after univariate analyses were assessed for multivariate Cox regression model using backward elimination. Variables were removed from the model if p > 0.1. Multivariate Cox analyses were performed to identify the significant parameters, and the constants and the standardized coefficients of the prognostic model. Based on this model, a nomogram was built to estimate the individual OS probability at 24 months.

Accepted manuscript

Two types of validation of the prognostic model were performed. An internal validation of the multivariate Cox model was first performed using the bootstrap method (1000 datasets constructed by random re-sampling with replacement from the original) [25]. This bootstrap method was used to estimate the adjusted c-index and the 95% confidence interval (95% CI) of each parameter. An internal calibration was performed to estimate the accuracy of the model. In the second step, the Cox model was used to compute the prognostic index for the patients of the validation cohort, and the corresponding c-index was computed. The individual probabilities computed by the model were averaged to produce the predicted survival curve. Three prognostic groups were created by categorizing the prognostic index computed from the model at the 50th and 84th percentiles. These groups were called low-, intermediate- and high-risk groups. The same cut-off was applied in the validation cohort.

All analyses were performed using R software 3.4.0 (R Development CoreTeam; http://www.r-project.org).

Results

The median values of different PET parameters (SUV_{Max}, MTV, TLG) for the training and the validation cohort are presented in Table S1.

Identification of the Cox model and nomogram to predict OS in the training cohort

The results from the univariate analysis are presented in Table 2. The prognostic value of survival of the different thresholds for the MTV, SUV mean, and TLG to predict OS is presented in Fig. 1. For the primary tumour, MTV with relative thresholds from 6 to 83% of SUV_{Max} and ab- solute thresholds from 1.5 to 6.5 was significantly correlated with OS, with a c-index ranging from 0.58 to 0.62 and from 0.57 to 0.6 respectively. For the lymph nodes, MTV with relative thresholds from 6 to 100% and absolute thresholds from 0 to 20 was significantly correlated with OS, with a c- index ranging from 0.54 to 0.61 and from 0.50 to 0.61 respectively. No interactions between tumour site or p16 status and PET parameters were observed. Due to the ab- sence of a difference in OS between hypopharynx and oro- pharynx cancer, these two tumour sites were combined into one group named pharynx cancer. Similarly, given that no significant difference was identified between negative and unknown p16 status, this parameter was recoded as positive versus negative/unknown status.

The MTVof the tumour with a relative threshold of 35% of the SUV_{Max}, the MTV of the lymph node with a relative threshold of 44% of the SUV_{Max} were the PET parameters achieving the highest c-index. In addition to these PET parameters, p16 status and N-classification were significant after univariate analysis and were assessed for multivariate analysis (Table 2).

The retained significant parameters from the multivariate analysis were the MTV of the tumour with a relative threshold of 35% of the SUV_{Max}, the MTV of the lymph node with a relative threshold of 44% of the SUV_{Max}, the tumour site, and p16 status (Table 3). The c-index of the model was 0.72 ($p < 0.001$). The β-coefficients of the corresponding Cox model are presented in Table 3, allowing the calculation of a prognostic index (OS probability) for each patient. Based on the Cox model, a nomogram was computed (Fig. 2).

Internal and external validations of the prognostic model

The adjusted c-index estimated from internal bootstrap validation of the Cox model was 0.71 ($p < 0.001$). The 95% CI values for the coefficient of the parameters of the model are presented in Table 3. Internal calibration revealed a very good adjustment between the predicted and observed OS at 24 months (Fig. S1).

The β-coefficients from the training model were applied to the external validation cohort, achieving a c-index of 0.66 ($p < 0.001$). The comparison between the predicted and observed OS is presented in Fig. 3. The prognostic index was computed for each patient. Based on the cut-off computed from the training cohort (Fig. S2A), three prognostic groups (high risk, intermediate risk, and low risk) were created for the validation cohort. OS differed significantly in these three risk groups (Fig. S2B).

Discussion

Accepted Manuscript

In this multicentre study including a large number of locally advanced head and neck cancer patients, MTV-T and MTV-N as continuous variables combined with tumour localization and p16 status were major predictors of the risk of death in LAHNC. Notably, the classical clinical variables [T and N classification, GTV (T and N), age, gender] were less or not significantly prognostic of OS (c-index ≤ 0.56). Furthermore, we proposed and successfully validated a prognostic model of survival, achieving a relatively high prediction capability (c- index = 0.72). This model may be used via our nomogram to estimate the individual survival of an external patient. The PET parameters can be easily computed by using the online free platform QuantImage [21].

Among PET parameters, MTV correlates with OS and disease-free survival (DFS) with a prognostic value clearly greater than the SUV_{Max} and less than the TLG [26, 27]. However, the majority of studies supporting MTV as a prognostic factor are retrospective or single-centre studies and/or lack external validation. Moreover, no prognostic model has been proposed. Only three studies performed external validation of the prognostic value of MTV and TLG [28–30], albeit with some major issues, thus limiting the impact of these studies. Indeed, TLG (using a specific cut-off of 58.7 ml) correlated with OS in a limited cohort of 52 patients [28], and a validation in a second cohort of 37 patients was per- formed after an adjustment of the TLG cut-off value of 141 ml [31]. Complete external validation requires assessing the performance of a predefined model in new data without refitting the variables of the model [32]. External validation was performed in another study including 168 patients (85 in the training cohort and 83 in the validation cohort), demonstrating that a 17-ml increase of the MTV was associated with a twofold increase in the risk of death [30]. However, this specific MTV value was based on a monocentric population-dependent characteristic (from the 25th to 75th percentile). Another study including 122 patients with oropharyngeal cancer from two different cohorts successfully per- formed external validation. Using two different thresholds for the primary tumour (35% of SUV_{Max}) and the lymph nodes (44% of the SUV_{Max}), MTV was correlated with OS [29]. However, this study exclusively included patients with oropharyngeal cancer. In the present study, we demonstrate that a single model with the same parameters was prognostic of OS for hypopharyngeal, oropharyngeal, and oral cavity cancers without an interaction between the tumour site and the MTV.

The reproducibility of the MTV or TLG is potentially limited by the initial definitions of these parameters, which are based on a threshold of SUV, either absolute (all pixel with SUV value > threshold) or relative (all pixel with SUV value > threshold% of SUV_{Max}). SUV may vary with PET scanner, certainly among scanners from different generations. However the use of relative threshold of SUV would partly avoid such difficulties for larger lesions, as SUV_{Max} in such regions would not be affected by TOF and PSF recovery correction. Most of the studies tested only one threshold (2.5 or 3, or 40 to 50%). Six studies compared only three or four different thresholds of MTV and/or TLG, most often using the same threshold of 40 and 50% of SUV_{Max} or 2.5 and 3 of absolute SUV [11–15, 33, 34]. However, the use of different thresholds within a reasonable range seems to have no major impact on the prognostic value of PET parameters. Indeed, a wide range of thresholds (from 0 to 20 mg/ml and from 0 to 100% of SUV_{Max}) was analysed for 123 patients with an oropharyngeal cancer. MTVs (of both primary tumour and lymph nodes) between 5 and 7 or between 30 and 64% were significantly correlated with DFS [35]. In our study, exploring continuous thresholds from 0 to 100% and from 0 to 20, MTV of the primary tumour computed with a relative threshold of 35% was the highest prognostic parameter of OS, with a c-index of 0.62. However, MTV of the primary tumour between 35 and 55% was also significantly correlated with OS, with slightly different of c-indexes (ranging from 0.6 to 0.62) (Fig. 1).

Most studies testing the prognostic value of PET imaging have exclusively focused on the primary tumour [16, 17, 36–39] because lymph node analysis appears particularly complex, given the number of lymph nodes and the choice of SUV_{Max} to compute MTV and/or TLG. A limited number of studies have computed MTV of the lymph nodes using an absolute threshold of 1.5 [40] or 2.5 [41] or a relative thresh- old of 50% of the SUV_{Max} only for the largest lymph node [14]. Based on PET performed during the third week of treatment for 75 patients with node-positive LAHNC, a reduction in the total node TLG of greater than 50% was highly correlated with an improvement in OS [41]. To compute lymph node MTV, we used SUV_{Max} of the primary tumour as the reference considering all the metastatic lymph nodes, which were defined based on the RECIST 1.1 criteria [42] and SUV_{Max}. In univariate analysis, we identified a large number of lymph node PET parameters prognostic of OS (Fig. 1), with c-indexes ranging from 0.5 to 0.61. Both analyses of the primary tumour and the lymph nodes were clearly justified, improving the prediction capability in multivariate analysis by up to 0.72. Moreover, the best clinical model (without PET parameters), including N-staging, p16 status and tumour site, achieved a c-index of 0.66, lower than the model including PET parameters. Furthermore, the external validation of this clinical model failed (c-index = 0.63, p = 0.06). In a recent study, a nomogram prognostic of overall survival for oropharyngeal cancer based on clinical and biological values

Accepted Manuscript

achieved a c-index of 0.68 in the validation cohort. However, this nomogram was for oropharynx cancer and included nine parameters [1]. Our model achieved a similar value of c-index (0.66) in the validation cohort with a fewer number of parameters (4), and could be used regardless the tumour site. The combination of PET and biological parameters may improve the prediction capabilities of such models.

Our study had some limitations. First, the analysis was retrospective, but the endpoint was OS. A large number of parameters (> 500) were explored, which may result in overfitting. However, the bootstrap method did not show ma- jor overfitting. Moreover, as the model is reproducible in an- other cohort, overfitting is clearly not an issue. P16-HPV sta- tus was missing for 40% of the patients and positive for 37% of the analysed patients. The prevalence of HPV in oropharyngeal cancer in Europe ranges from 23% [43] to 70% [44], with large differences noted between countries. Similarly to widely published data [45], we found that positive p16 status was associated with better survival, and p16 was retained in the multivariate analysis. P16-positive tumours were previously associated with a higher MTV than for p16-negative tumours [46]. We found a similar trend in our study (data not shown); however, it was not significant. This lack of interaction between p16 status and MTV may be explained by missing data with regard to p16. Important differences in age, use of cetuximab, T-classification, p16 status, tumor and nodal volume (GTV), and PET/CT scanner between the training and the validation cohorts were observed. Despite these differences, the very good prognostic performance obtained with the training cohort was confirmed for the validation cohort. The good prognostic value of our PET-based parameters, in- dependently of these differences, could be considered a strength of the nomogram in a non-selected population. Another issue concerns N3 stage patients for whom ^{18}F- FDG -based MTV may not take into account hypoxic areas known to be more radioresistant [47]. Indeed, other hypoxia radiotracers tested in the head and neck, such as ^{18}F- fluoromisonidazole (FMISO) and ^{18}F-fluoroazomycin arabinoside (FAZA), potentially correlate with OS [48–50]. Moreover, we did not perform more complex analyses, such as adaptive thresholds [31, 51] and gradient-based methods [52], which may more accurately define tumour volume [53]. The reconstruction algorithms used were not EARL- compliant [54]. It would be expected that EARL-compliant data would enhance the significance and transposability of the results. We only analysed pre-treatment PET imaging, and it is possible that combining both pre-treatment and early pre-treatment imaging may further improve the capability of predicting patient outcome [55–57]. Finally, we exclusively investigated PET imaging. However, CT scan [58] and MRI image [59] analyses are also prognostic of patient outcome based on radiomic analysis including a large number of parameters.

Conclusion

In LAHNC treated with CRT or RT + cetuximab, the combi- nation of PET parameters from both the primary tumour and lymph nodes with tumour site and p16 status appears particularly useful for predicting OS. We provided and validated a prognostic model of OS, and proposed a nomogram to predict OS for an external patient. Indeed, patients with worse out- comes could be good candidates for treatment intensification, such as dose escalation or combining new systemic treatment.

Funding This work was partly supported by the Swiss National Science Foundation with grant agreement PZ00P2_154891 and 205320_179069 (A. Depeursinge).

Compliance with ethical standards

Conflict of interest All authors declare no conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institu- tional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol. 2010;17:1471–4. https://doi.org/10.1245/ s10434-010-0985-4.
2. Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-analysis of chemotherapy on head and neck cancer. Lancet. 2000;355:949–55.

Accepted Manuscript

3. Bernier J, Domenge C, Ozsahin M, Matuszewska K, Lefebvre JL, Greiner RH, et al. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. N Engl J Med. 2004;350:1945–52. https://doi.org/10.1056/NEJMoa032641.
4. Bourhis J, Sire C, Graff P, Gregoire V, Maingon P, Calais G, et al. Concomitant chemoradiotherapy versus acceleration of radiothera- py with or without concomitant chemotherapy in locally advanced head and neck carcinoma (GORTEC 99-02): an open-label phase 3 randomised trial. Lancet Oncol. 2012;13:145–53. https://doi.org/ 10.1016/s1470-2045(11)70346-1.
5. Bonner JA, Harari PM, Giralt J, Cohen RB, Jones CU, Sur RK, et al. Radiotherapy plus cetuximab for locoregionally advanced head and neck cancer: 5-year survival data from a phase 3 randomised trial, and relation between cetuximab-induced rash and survival. Lancet Oncol. 2010;11:21–8. https://doi.org/10.1016/s1470- 2045(09)70311-0.
6. Chajon E, Lafond C, Louvel G, Castelli J, Williaume D, Henry O, et al. Salivary gland-sparing other than parotid-sparing in definitive head-and-neck intensity-modulated radiotherapy does not seem to jeopardize local control. Radiat Oncol. 2013;8:132. https://doi.org/ 10.1186/1748-717x-8-132.
7. Schwartz DL, Harris J, Yao M, Rosenthal DI, Opanowski A, Levering A, et al. Metabolic tumor volume as a prognostic imaging-based biomarker for head-and-neck cancer: pilot results from radiation therapy oncology group protocol 0522. Int J Radiat Oncol Biol Phys. 2015;91:721–9. https://doi.org/10.1016/j. ijrobp.2014.12.023.
8. Moon SH, Choi JY, Lee HJ, Son YI, Baek CH, Ahn YC, et al. Prognostic value of volume-based positron emission tomography/computed tomography in patients with nasopharyngeal carcinoma treated with concurrent chemoradiotherapy. Clin Exp Otorhinolaryngol. 2015;8:142–8. https://doi.org/10.3342/ceo. 2015.8.2.142.

9. Cacicedo J, Navarro A, Del Hoyo O, Gomez-Iturriaga A, Alongi F, Medina JA, et al. Role of [18F] fluorodeoxyglucose PET/CT in head and neck oncology: the point of view of the radiation oncol- ogist. Br J Radiol. 2016;89:20160217. https://doi.org/10.1259/bjr. 20160217.
10. Castelli J, Depeursinge A, de Bari B, Devillers A, de Crevoisier R, Bourhis J, et al. Metabolic tumor volume and total lesion glycolysis in oropharyngeal cancer treated with definitive radiotherapy: which threshold is the best predictor of local control? Clin Nucl Med. 2017;42:e281–e5. https://doi.org/10.1097/rlu.0000000000001614.
11. Schinagl DA, Span PN, Oyen WJ, Kaanders JH. Can FDG PET predict radiation treatment outcome in head and neck cancer? Results of a prospective study. Eur J Nucl Med Mol Imaging. 2011;38:1449–58. https://doi.org/10.1007/s00259-011-1789-x.
12. Kao CH, Lin SC, Hsieh TC, Yen KY, Yang SN, Wang YC, et al. Use of pretreatment metabolic tumour volumes to predict the outcome of pharyngeal cancer treated by definitive radiotherapy. Eur J Nucl Med Mol Imaging. 2012;39:1297–305. https://doi.org/10. 1007/s00259-012-2127-7.
13. Cheng NM, Fang YH, Lee LY, Chang JT, Tsan DL, Ng SH, et al. Zone-size nonuniformity of 18F-FDG PET regional textural features predicts survival in patients with oropharyngeal cancer. Eur J Nucl Med Mol Imaging. 2015;42:419–28. https://doi.org/10. 1007/s00259-014-2933-1.
14. Lin YC, Chen SW, Hsieh TC, Yen KY, Yang SN, Wang YC, et al. Risk stratification of metastatic neck nodes by CT and PET in patients with head and neck cancer receiving definitive radiotherapy. J Nucl Med. 2015;56:183–9. https://doi.org/10.2967/jnumed.114. 148023.
15. Yabuki K, Shiono O, Komatsu M, Sano D, Nishimura G, Takahashi M, et al. Predictive and prognostic value of metabolic tumor vol- ume (MTV) in patients with laryngeal carcinoma treated by radio- therapy (RT) /concurrent chemoradiotherapy (CCRT). PLoS One. 2015;10:e0117924. https://doi.org/10.1371/journal.pone.0117924.
16. Abgral R, Keromnes N, Robin P, Le Roux PY, Bourhis D, Palard X, et al. Prognostic value of volumetric parameters measured by 18F- FDG PET/CT in patients with head and neck squamous cell carci- noma. Eur J Nucl Med Mol Imaging. 2014;41:659–67. https://doi. org/10.1007/s00259-013-2618-1.
17. Abgral R, Valette G, Robin P, Rousset J, Keromnes N, Le Roux PY, et al. Prognostic evaluation of percentage variation of metabolic tumor burden calculated by dual-phase (18) FDG PET-CT imaging in patients with head and neck cancer. Head Neck. 2016; 38(Suppl 1:E600-6. https://doi.org/10.1002/hed.24048.
18. Park GC, Kim JS, Roh JL, Choi SH, Nam SY, Kim SY. Prognostic value of metabolic tumor volume measured by 18F-FDG PET/CT in advanced-stage squamous cell carcinoma of the larynx and hy- popharynx. Ann Oncol. 2013;24:208–14. https://doi.org/10.1093/ annonc/mds247.
19. Mohan R, Wu Q, Manning M, Schmidt-Ullrich R. Radiobiological considerations in the design of fractionation strategies for intensity- modulated radiation therapy of head and neck cancers. Int J Radiat Oncol Biol Phys. 2000;46:619–30.
20. Lee N, Chuang C, Quivey JM, Phillips TL, Akazawa P, Verhey LJ, et al. Skin toxicity due to intensity-modulated radiotherapy for head-and-neck carcinoma. Int J Radiat Oncol Biol Phys. 2002;53: 630–7.
21. Dicente Cid Y, Castelli J, Schaer R, Scher N, Pomoni A, Prior JO, et al. Chapter 12 — Quantimage: an online tool for high-throughput 3D radiomics feature extraction in PET-CT. In: Depeursinge A, Al- Kadi OS, Mitchell JR (Eds) Biomedical texture analysis: London, Academic; 2017. pp 349-77.
22. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials. 1996;17:343–6.
23. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Br J Cancer. 2015;112:251–9. https://doi.org/10.1038/bjc.2014.639.
24. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996;15:361–87. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361:: AID-SIM168>3.0.CO;2-4.
25. Efron B, Gong G. A leisurely look at the bootstrap, the jackknife, and cross-validation. Am Stat. 1983;37:36–48. https://doi.org/10. 2307/2685844.
26. Castelli J, De Bari B, Depeursinge A, Simon A, Devillers A, Roman Jimenez G, et al. Overview of the predictive value of quantitative 18 FDG PET in head and neck cancer treated with chemo- radiotherapy. Crit Rev Oncol Hematol. 2016;108:40–51. https:// doi.org/10.1016/j.critrevonc.2016.10.009.
27. Bonomo P, Merlotti A, Olmetto E, Bianchi A, Desideri I, Bacigalupo A, et al. What is the prognostic impact of FDG PET in locally advanced head and neck squamous cell carcinoma treated with concomitant chemo-radiotherapy? A systematic review and meta-analysis. Eur J Nucl Med Mol Imaging. 2018. https://doi. org/10.1007/s00259-018-4065-5.
28. Apostolova I, Steffen IG, Wedel F, Lougovski A, Marnitz S, Derlin T, et al. Asphericity of pretherapeutic tumour FDG uptake provides independent prognostic value in head-and-neck cancer. Eur Radiol. 2014;24:2077–87. https://doi.org/10.1007/s00330-014-3269-8.
29. Castelli J, Depeursinge A, Ndoh V, Prior JO, Ozsahin M, Devillers A, et al. A PET-based nomogram for oropharyngeal cancers. Eur J Cancer. 2017;75:222–30. https://doi.org/10.1016/j.ejca.2017.01. 018.

Accepted Manuscript

30. Tang C, Murphy JD, Khong B, La TH, Kong C, Fischbein NJ, et al. Validation that metabolic tumor volume predicts outcome in head- and-neck cancer. Int J Radiat Oncol Biol Phys. 2012;83:1514–20. https://doi.org/10.1016/j.ijrobp.2011.10.023.
31. Hofheinz F, Lougovski A, Zophel K, Hentschel M, Steffen IG, Apostolova I, et al. Increased evidence for the prognostic value of primary tumor asphericity in pretherapeutic FDG PET for risk strat- ification in patients with head and neck cancer. Eur J Nucl Med Mol Imaging. 2015;42:429–37. https://doi.org/10.1007/s00259-014- 2953-x.
32. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol. 2013;13:33. https://doi.org/10.1186/1471-2288-13-33.
33. Murphy JD, La TH, Chu K, Quon A, Fischbein NJ, Maxim PG, et al. Postradiation metabolic tumor volume predicts outcome in head- and-neck cancer. Int J Radiat Oncol Biol Phys. 2011;80:514–21. https://doi.org/10.1016/j.ijrobp.2010.01.057.
34. Moon SH, Choi JY, Lee HJ, Son YI, Baek CH, Ahn YC, et al. Prognostic value of 18F-FDG PET/CT in patients with squamous cell carcinoma of the tonsil: comparisons of volume-based meta- bolic parameters. Head Neck. 2013;35:15–22. https://doi.org/10.1002/hed.22904.
35. Castelli J, Depeursinge A, de Bari B, Devillers A, de Crevoisier R, Bourhis J, et al. Metabolic tumor volume and total lesion glycolysis in oropharyngeal cancer treated with definitive radiotherapy: which threshold is the best predictor of local control? Clin Nucl Med. 2017;42(6):e281–e285. https://doi.org/10.1097/RLU. 0000000000001614.
36. Akagunduz OO, Savas R, Yalman D, Kocacelebi K, Esassolak M. Can adaptive threshold-based metabolic tumor volume (MTV) and lean body mass corrected standard uptake value (SUL) predict prognosis in head and neck cancer patients treated with definitive radiotherapy/chemoradiotherapy? Nucl Med Biol. 2015;42:899– 904. https://doi.org/10.1016/j.nucmedbio.2015.06.007.
37. Cheng NM, Fang YH, Chang JT, Huang CG, Tsan DL, Ng SH, et al. Textural features of pretreatment 18F-FDG PET/CT images: prognostic significance in patients with advanced T-stage oropha- ryngeal squamous cell carcinoma. J Nucl Med. 2013;54:1703–9. https://doi.org/10.2967/jnumed.112.119289.
38. Hanamoto A, Tatsumi M, Takenaka Y, Hamasaki T, Yasui T, Nakahara S, et al. Volumetric PET/CT parameters predict local response of head and neck squamous cell carcinoma to chemora- diotherapy. Cancer Med. 2014;3:1368–76. https://doi.org/10.1002/ cam4.295.
39. Hentschel M, Appold S, Schreiber A, Abolmaali N, Abramyuk A, Dorr W, et al. Early FDG PET at 10 or 20 Gy under chemoradiotherapy is prognostic for locoregional control and overall survival in patients with head and neck cancer. Eur J Nucl Med Mol Imaging. 2011;38:1203–11. https://doi.org/10.1007/s00259-011-1759-3.
40. Alluri KC, Tahari AK, Wahl RL, Koch W, Chung CH, Subramaniam RM. Prognostic value of FDG PET metabolic tumor volume in human papillomavirus-positive stage III and IV oropha- ryngeal squamous cell carcinoma. AJR Am J Roentgenol. 2014;203:897–903. https://doi.org/10.2214/AJR.14.12497.
41. Lin P, Min M, Lee M, Holloway L, Forstner D, Bray V, et al. Nodal parameters of FDG PET/CT performed during radiotherapy for locally advanced mucosal primary head and neck squamous cell carcinoma can predict treatment outcomes: SUVmean and response rate are useful imaging biomarkers. Eur J Nucl Med Mol Imaging. 2017;44(5):801-811 https://doi.org/10.1007/s00259-016-3584-1.
42. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45: 228–47. https://doi.org/10.1016/j.ejca.2008.10.026.
43. Kreimer AR, Clifford GM, Boyle P, Franceschi S. Human papillo- mavirus types in head and neck squamous cell carcinomas world- wide: a systematic review. Cancer Epidemiol Biomark Prev. 2005;14:467–75. https://doi.org/10.1158/1055-9965.EPI-04-0551.
44. Mehanna H, Beech T, Nicholson T, El-Hariry I, McConkey C, Paleri V, et al. Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer–systematic re- view and meta-analysis of trends by time and region. Head Neck. 2013;35:747–55. https://doi.org/10.1002/hed.22015.
45. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, et al. Human papillomavirus and survival of patients with oro- pharyngeal cancer. N Engl J Med. 2010;363:24–35. https://doi.org/ 10.1056/NEJMoa0912217.
46. Kendi AT, Magliocca K, Corey A, Nickleach DC, Galt J, Higgins K, et al. Do 18F-FDG PET/CT parameters in oropharyngeal and oral cavity squamous cell carcinomas indicate HPV status? Clin Nucl Med. 2015;40:e196–200. https://doi.org/10.1097/rlu.0000000000000691.
47. Brizel DM, Sibley GS, Prosnitz LR, Scher RL, Dewhirst MW. Tumor hypoxia adversely affects the prognosis of carcinoma of the head and neck. Int J Radiat Oncol Biol Phys. 1997;38:285–9.
48. Graves EE, Hicks RJ, Binns D, Bressel M, Le Q-T, Peters L, et al. Quantitative and qualitative analysis of [(18)F] FDG and [(18)F] FAZA positron emission tomography of head and neck cancers and associations with HPV status and treatment outcome. Eur J Nucl Med Mol Imaging. 2016;43:617–25. https://doi.org/10.1007/ s00259-015-3247-7.
49. Beck R, Roper B, Carlsen JM, Huisman MC, Lebschi JA, Andratschke N, et al. Pretreatment 18F-FAZA PET predicts suc- cess of hypoxia-directed radiochemotherapy using tirapazamine. J Nucl Med. 2007;48:973–80. https://doi.org/10.2967/jnumed.106.038570.
50. Thorwarth D, Eschmann SM, Holzner F, Paulsen F, Alber M. Combined uptake of [18F] FDG and [18F] FMISO correlates with radiation therapy outcome in head-and-neck cancer patients. Radiother Oncol. 2006;80:151–6. https://doi.org/10.1016/j.radonc. 2006.07.033.
51. Kikuchi M, Koyasu S, Shinohara S, Usami Y, Imai Y, Hino M, et al. Prognostic value of pretreatment 18F-fluorodeoxyglucose positron emission tomography/CT volume-based parameters in patients with oropharyngeal squamous cell carcinoma with known p16 and p53 status. Head Neck. 2015;37:1524–31. https://doi.org/10.1002/hed. 23784.
52. Romesser PB, Qureshi MM, Shah BA, Chatburn LT, Jalisi S, Devaiah AK, et al. Superior prognostic utility of gross and meta- bolic tumor volume compared to standardized uptake value using PET/CT in head and neck squamous cell carcinoma patients treated with intensity-modulated radiotherapy. Ann Nucl Med. 2012;26: 527–34. https://doi.org/10.1007/s12149-012-0604-5.
53. Arens AI, Troost EG, Hoeben BA, Grootjans W, Lee JA, Gregoire V, et al. Semiautomatic methods for segmentation of the prolifera- tive tumour volume on sequential FLT PET/CT images in head and neck carcinomas and their relation to clinical outcome. Eur J Nucl Med Mol Imaging. 2014;41:915–24. https://doi.org/10.1007/ s00259-013-2651-0.
54. Kaalep A, Sera T, Oyen W, Krause BJ, Chiti A, Liu Y, et al. EANM/ EARL FDG-PET/CT accreditation — summary results from the first 200 accredited imaging systems. Eur J Nucl Med Mol Imaging. 2018;45:412–22. https://doi.org/10.1007/s00259-017-3853-7.
55. Brun E, Kjellen E, Tennvall J, Ohlsson T, Sandell A, Perfekt R, et al. FDG PET studies during treatment: prediction of therapy outcome in head and neck squamous cell carcinoma. Head Neck. 2002;24:127–35.

Accepted Manuscript

56. Chen SW, Hsieh TC, Yen KY, Yang SN, Wang YC, Chien CR, et al. Interim FDG PET/CT for predicting the outcome in patients with head and neck cancer. Laryngoscope. 2014;124:2732–8. https://doi. org/10.1002/lary.24826.
57. Min M, Lin P, Lee MT, Shon IH, Lin M, Forstner D, et al. Prognostic role of metabolic parameters of (18)F-FDG PET-CT scan performed during radiation therapy in locally advanced head and neck squamous cell carcinoma. Eur J Nucl Med Mol Imaging. 2015;42:1984–94. https://doi.org/10.1007/s00259-015-3104-8.
58. Leijenaar RT, Carvalho S, Hoebers FJ, Aerts HJ, van Elmpt WJ, Huang SH, et al. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. Acta Oncol. 2015;54:1423–9. https://doi.org/10.3109/0284186x. 2015.1061214.
59. Min M, Lin P, Liney G, Lee M, Forstner D, Fowler A, et al. A review of the predictive role of functional imaging in patients with mucosal primary head and neck cancer treated with radiation ther- apy. J Med Imaging Radiat Oncol. 2017;61:99–123. https://doi.org/ 10.1111/1754-9485.12496.

Accepted Manuscript

Table 1 Patient, tumour, and treatment characteristics and follow-up and treatment results for the training and validation cohorts

Characteristics	Training cohort (one centre) ($N = 127$ pts)	Validation cohort (three centres) ($N = 110$ pts)	P value*
Patients and tumours			
Mean age, years (SD)	59 (8.9)	62 (8.4)	< 0.01
Gender, *n* (%)			0.9
Male	107 (85%)	92 (83.6%)	
Female	20 (15%)	18 (16.4%)	
Smoking status, *n* (%)			0.6
Yes	117 (86.6%)	92 (83.6%)	
No	10 (13.4%)	18 (16.4%)	
Tumour site, *n* (%)			0.3
Oropharynx	89 (70.1%)	74 (67.3%)	
Oral cavity	14 (11%)	8 (7.2%)	
Hypopharynx	24 (18.9%)	28 (25.5%)	
T classification, *n* (%)			0.6
T1	4 (3.2%)	5 (4.5%)	
T2	29 (22.8%)	27 (24.5%)	
T3	49 (38.6%)	48 (43.6%)	
T4	45 (35.4%)	30 (27.3%)	
N classification, *n* (%)			0.3
N0	19 (15%)	20 (18.2%)	
N1	17 (13.4%)	19 (17.3%)	
N2	83 (65.3%)	69 (62.3%)	
N3	8 (6.3%)	2 (1.8%)	
Stage (AJCC)			
III	18 (14.2%)	36 (32.8%)	0.03
IV	109 (85.8%)	74 (67.2%)	
Mean tumour volume ** in ml (SD)	38.9 (40.9)	20.5 (21.8)	< 0.01
p16 status, *n* (%)			0.3
Positive	17 (13.4%)	21 (19.1%)	
Negative	53 (41.7%)	49 (44.5%)	
Unknown	57 (44.9%)	40 (36.4%)	
Concomitant systemic treatment with RT			
N (%)			< 0.01
Cisplatin [3]	86 (67.7%)	60 (54.5%)	
Carboplatin – 5FU [4]	11 (8.7%)	19 (17.3%)	
Cetuximab [5]	30 (23.6%)	30 (27.2%)	
Treatment results			
Follow-up (months) (min–max)	45 (6–92)	23 (3–57)	< 0.01
2-year LRC rate [95% CI]	60% [52–70%]	79% [71–88%]	< 0.01
2-year OS rate [95% CI]	60% [52–69%]	73% [64–83%]	0.02

SD = standard deviation, 95% CI = 95% confidence interval

* = Mann-Whitney test for continuous variables, chi^2 for dichotomous variables

**= Delineated on the planning CT (GTV), including both primary tumour and lymph nodes

Accepted Manuscript

Table 2 Univariate analysis for OS in the training cohort

Parameters	HR [95% CI]	C-index	*P* value
Clinical parameters			
Age (in years)	–	–	0.9
Gender			
Male	–	–	0.1
Female			
Smoking status, *n* (%)			
Yes	–	–	0.1
No			
GTV (in ml)	–	–	0.07
T classification			
T1-2			
T3	–	–	0.16
T4			
N classification			
N0-1	1	0.56	0.02
N2	1.4 [0.8–2.3]		
N3	4.1 [1.7–10]		
Tumour site			
Oropharynx	1	0.58	<0.01
Hypopharynx	1.4 [0.7–2.6]		
Oral cavity	3.2 [1.7–6]		
p16 status			
Positive	1	0.56	<0.01
Negative/unknown	5.7 [1.4–23.4]		
Concomitant systemic treatment with RT			
Cisplatin			
Carboplatin – 5FU	–	–	0.08
Cetuximab			
PET Parameters (highest c-index only with p value <0.05)			
MTV-T (per 10 ml)			
Absolute threshold (SUV = 3.5)	1.11 [1.02–1.23]	0.6	0.03
Relative threshold (35% of SUV_{Max})	1.25 [1.09–1.43]	0.62	0.001
MTV-N (per 10 ml)			
Absolute threshold (SUV = 3)	1.15 [1.07–1.24]	0.61	<0.001
Relative threshold (44% of SUV_{Max})	1.5 [1.2–1.8]	0.61	<0.001
TLG T (per 10 ml)			
Absolute threshold (SUV = 3)	1.01 [1–1.02]	0.59	0.05
Relative threshold (40% of SUV_{Max})	1.02 [1–1.03]	0.61	0.02
TLG N (per 10 ml)			
Absolute threshold (SUV = 3)	1.02 [1.01–1.04]	0.61	< 0.001
Relative threshold (46% of SUV_{Max})	1.05 [1.03–1.08]	0.60	< 0.001
SUV Mean T			
Absolute threshold (SUV = 1.5)	1.24 [1.01–1.51]	0.57	0.04
SUV Mean *N*			
Absolute threshold (SUV = 5.5)	1.08 [1.00–1.15]	0.61	0.03
Relative threshold (78% of SUV_{Max})	1.09 [1.05–1.14]	0.62	< 0.001

For PET parameters, data are only provided for absolute and relative thresholds with the highest c-index and $p < 0.05$. All the significant PET parameters are presented in Fig. 1. HR = hazard ratio, CI = confidence interval, GTV = gross tumour volume, SUV = standard uptake value, MTV = metabolic tumour volume, TLG = total lesion glycolysis

Accepted Manuscript

Table 3 Significant predictors of OS in multivariate analysis in the training cohort

Parameters	Multivariate Cox analysis						Cox model bootstrap validation (1000 samples)		
	HR [95% CI]	*P*	c-index	Standardized regression coefficient	SE	H_0	HR [95% CI]	*P*	c-index
Tumour site (pharynx or oral cavity)	2.8 [1.5–5.3]	< 0.01	0.72	1.02	0.32	−2.04	2.8 [1.4– 6.3]	< 0.01	0.71
P16 status (positive or negative/unknown)	4.8 [1.2–19.8]	0.03		1.57	0.72		4.8 [1.4–1530]	0.03	
MTV_T_35% (per 10 ml)	1.18 [1.02–1.37]	0.02		0.27	0.11		1.18 [1– 1.4]	0.05	
MTV_N_44% (per 10 ml)	1.58 [1.3– 1.93]	< 0.01		0.27	0.05		1.57 [1.3– 1.91]	< 0.01	

H_0 = baseline hazard, HR = hazard ratio, CI = confidence interval, SE = standard error, MTV_T_35 = metabolic tumour volume of the tumour computed with a relative threshold at 35% of SUV_{Max}, MTV_N_44 = metabolic tumour volume of the lymph node computed with a relative threshold at 44% of SUV_{Max}. Tumour site = oropharynx/hypopharynx or oral cavity

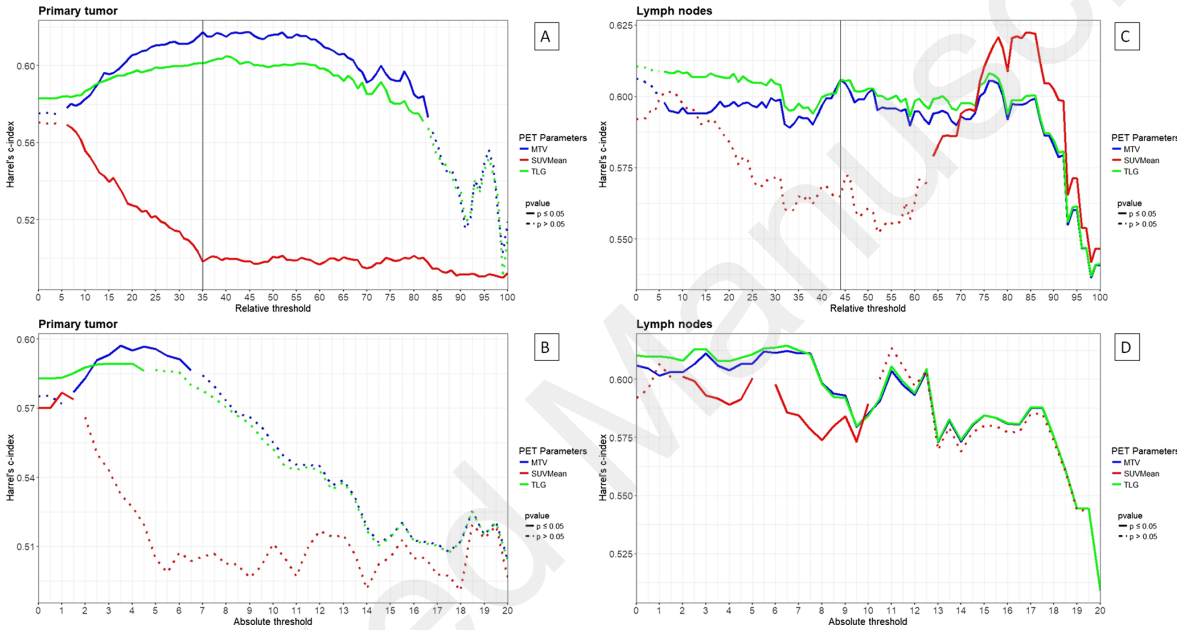


Fig. 1 C-index values for metabolic tumour volume (MTV), SUV_{Mean} and total lesion glycolysis (TLG) of the primary tumour (**a** and **b**) and the lymph nodes (**c** and **d**) computed with different relative thresholds [from 0 to 100% of SUV_{Max} (**a** and **c**) and from 0 to 20 (**b** and **d**)] to predict OS. *Full line* corresponds to a *p* value ≤ 0.05. MTV of the primary tumour computed with a relative threshold between 35 and 55% of the SUV_{Max} reached the highest c-index. *Vertical black line* corresponds to the optimum threshold for the MTV (highest c-index). The zero-threshold value corresponds to the region of interest (ROI) computed by adding a 3D margin of 5 mm to GTV-T and GTV-N

Accepted Manuscript

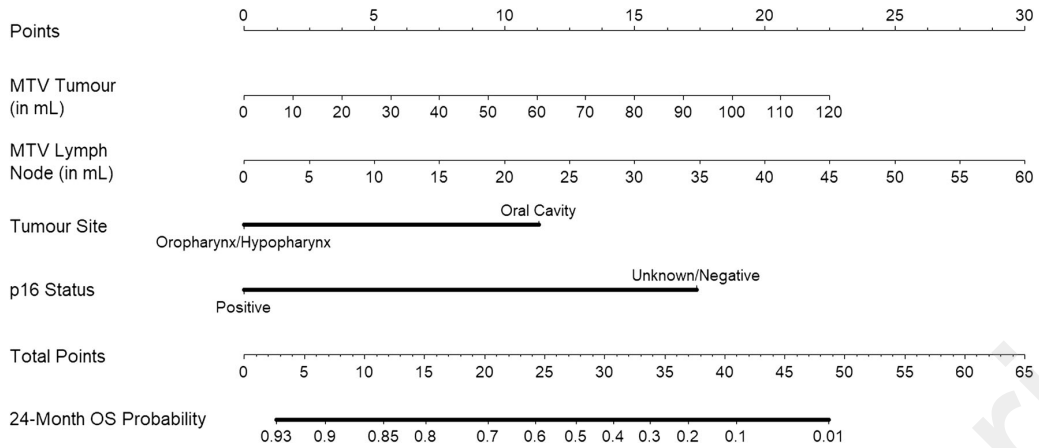


Fig. 2 Nomogram to predict 24-month OS. For each parameter, the corresponding points are obtained by drawing a line upward from the corresponding values to the *Points* line. The total points for each patient are obtained by summing the points for each of the individual factors in the nomogram, and this value is plotted on the *Total Points* line. A line is drawn down to determine the corresponding predictions of 24-month OS. MTV_N = metabolic tumour volume of the lymph node computed with a threshold = 44% of the SUV_{Max}. MTV_T = metabolic tumour volume of the tumour computed with a threshold = 35% of the SUV_{Max}. For example, an oropharynx cancer lesion that is p16 negative with an MTV tumour of 20 ml and an MTV lymph node of 5 ml corresponds to a 24-month OS probability of 60% (oropharynx cancer = 0 pts., p16 negative = 17.5 pts., MTV tumour of 20 ml = 4.25 pts., MTV lymph node of 5 ml = 2.5 pts. Total points = 24.25)

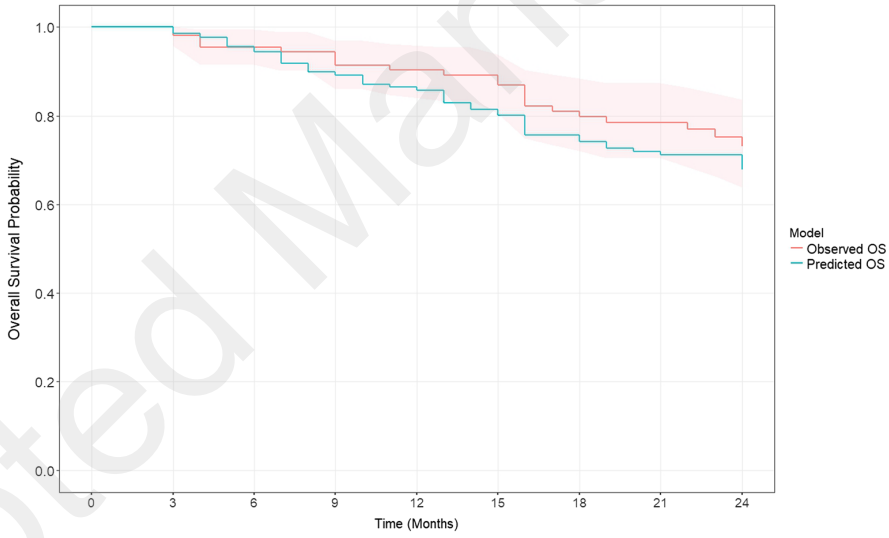


Fig. 3 Predicted and observed Kaplan–Meier curves of OS for the validation cohort. The *red line* indicates the observed OS for the validation cohort, and the *red area* corresponds to the 95% confidence interval. The *green line* indicates the predicted OS when applying the prognostic model to the patients of the validation cohort. The predicted OS is included in the 95% CI of the observed OS

Accepted Manuscript