Metabolic Tumor Volume and Total Lesion Glycolysis in Oropharyngeal Cancer Treated With Definitive Radiotherapy

Which Threshold Is the Best Predictor of Local Control?

Joël Castelli, MD, MSc,*†‡‡ Adrien Depeursinge, PhD,§|| Berardino de Bari, MD,*¶
Anne Devillers, MD,** Renaud de Crevoisier, MD, PhD,†‡†‡ Jean Bourhis, MD, PhD,* and John O. Prior, MD, PhD‡‡‡

Purpose: In the context of oropharyngeal cancer treated with definitive radiotherapy, the aim of this retrospective study was to identify the best threshold value to compute metabolic tumor volume (MTV) and/or total lesion glycolysis to predict local-regional control (LRC) and disease-free survival.

Methods: One hundred twenty patients with a locally advanced oropharyngeal cancer from 2 different institutions treated with definitive radiotherapy underwent FDG PET/CT before treatment. Various MTVs and total lesion glycolysis were defined based on 2 segmentation methods: (i) an absolute threshold of SUV (0–20 g/mL) or (ii) a relative threshold for SUVmax (0%–100%). The parameters’ predictive capabilities for disease-free survival and LRC were assessed using the Harrell C-index and Cox regression model.

Results: Relative thresholds between 40% and 68% and absolute threshold between 5.5 and 7 had a similar predictive value for LRC (C-index = 0.65 and 0.64, respectively). Metabolic tumor volume had a higher predictive value than gross tumor volume (C-index = 0.61) and SUVmax (C-index = 0.54). Metabolic tumor volume computed with a relative threshold of 51% of SUVmax was the best predictor of disease-free survival (hazard ratio, 1.23 [per 10 mL], P = 0.009) and LRC (hazard ratio: 1.22 [per 10 mL], P = 0.02).

Conclusions: The use of different thresholds within a reasonable range (between 5.5 and 7 for an absolute threshold and between 40% and 68% for a relative threshold) seems to have no major impact on the predictive value of MTV. This parameter may be used to identify patient with a high risk of recurrence and who may benefit from treatment intensification.

Key Words: metabolic tumor volume, oropharyngeal cancer, PET, threshold

(Clin Nucl Med 2017;00: 00–00)

18F-FDG PET/CT allows to quantify the metabolic activity of a tumor (glycolysis) and has become a reference tool in oncology for staging, radiotherapy planning, and monitoring tumor response in many cancers.1,2 Compared with other diagnostic modalities, PET imaging allows a most accurate nodal staging of locally advanced head and neck cancer3–8 and could result in changing the therapeutic management in nearly 15% of patients.9 The SUVmax corresponds to the maximal pixel value in the tumor. Thanks to its ease of use and relative robustness, it is one of the most widely used parameters in clinical practice. However, SUVmax is not representative of nonhomogeneous overall tumor uptake. More recently, volumetric PET parameters—metabolic tumor volume (MTV) and total lesion glycolysis (TLG)—have been correlated with clinical outcome.5–8 Nonetheless, these parameters require a tumor segmentation that is classically defined by either a percentage of the SUVmax or absolute SUV as the lowest threshold for inclusion. The optimal SUV threshold for clinical outcome prediction in head and neck cancer is not well defined. Few studies have compared different thresholds of MTV and/or TLG9–14 and a large majority of studies using the same thresholds of 40% SUVmax15 or a fixed SUV threshold of greater than 2.5.16 In the context of oropharyngeal cancer treated with definitive radiotherapy, the aim of this retrospective study was to identify the best threshold value to compute MTV and/or TLG in order to predict clinical outcome.

MATERIALS AND METHODS

All consecutive patients from 1 cancer center and 1 university hospital treated with definitive concurrent chemoradiotherapy or radiotherapy–cetuximab for a locally advanced oropharyngeal carcinoma between January 2010 and December 2015 were retrospectively analyzed. The study enrolled a total of 122 patients. All tumors were locally advanced (stage III or IV, American Joint Committee on Cancer seventh edition).

Treatment and Planning

All patients underwent intensity-modulated radiotherapy using volumetric modulated arc therapy (Rennes) or helical tomotherapy (Lausanne). A total dose of 70 Gy (2 Gy/fraction per day, 35 fractions [Rennes]); or 2.12 Gy/fractions per day, 33 fractions [Lausanne], with a simultaneous integrated boost technique)17 was delivered combined to concomitant chemotheraphy18,19 or cetuximab20 if the patients were not fit for chemotherapy. The modality of planning and treatment were the same as previously described.21 The study was approved by both institutional ethical committees (NCT02469922).

PET/CT Acquisition

All patients underwent FDG PET/CT for staging before treatment. For (Rennes), the patients lasted at least 4 hours prior to injection of 4 MBq/kg of 18F-FDG (Flucis). Blood glucose levels were checked prior to the injection of 18F-FDG. If not contraindicated, intravenous contrast agents were administered before CT scanning. After a 60-minute uptake period of rest, patients were imaged with
For each patient, tumor gross tumor volume (GTV-T) and nodal GTV (GTV-N) were manually segmented on each PET/CT by the same radiation oncologist, experienced in head and neck cancer treatments. A region of interest (ROI) was computed by adding 3-dimensional margins of 10 mm to GTV-T and GTV-N (Fig. 1).

A set of quantitative parameters based on SUV histograms was extracted from ROI-T and ROI-N in PET images. SUVmax was first computed from ROI-T as the maximum SUV in the delineated volume. Several metabolic volumes were subsequently defined based on 2 segmentation methods: (i) an absolute threshold of SUV (ranging from 0 to 20 g/mL, 0.5-g/mL steps) or (ii) a relative threshold of SUVmax (0%–100%, 1% steps). Metabolic tumor volume was computed as the metabolic volume of the segmented region in milliliters and TLG as SUVmean × MTV of the corresponding thresholded region.

Statistical Analysis

Patients alive at the time of analysis were censored at the date of last follow-up. Disease-free survival (DFS) was calculated from the first day of radiotherapy (chemoradiotherapy) to the date of first event (local or distant recurrence or death). Locoregional control (LRC) was calculated from the first day of radiotherapy to the date of first recurrence in primary tumor and/or lymph node. Follow-up was calculated using a reverse Kaplan-Meier estimation. Disease-free survival and overall survival (OS) estimations were computed using the Kaplan-Meier method, and 2-sided log-rank test was used to compare groups.

The association of the PET pretreatment parameters with DFS and OS was assessed using univariate Cox analyses. Harrell C-index (C-index) was used to compare different models (C-index ≈ 0.5 → not predictive, C-index ≈ 1 → predictive). The C-index was used to determine the optimal SUV threshold giving the most predictive value for each PET parameter. Factors with significance of $P < 0.1$ and with the highest C-index after univariate analysis
were assessed for multivariate Cox regression model using backward elimination. Variables were removed from the model if \( P > 0.1 \).

Two prognostic risk groups were identified based on the estimated optimal cutoff point by the Hothorn and Lausen\(^{24}\) method. Kaplan-Meier method was used to evaluate this cutoff.

All analyses were performed using R software 3.2.4 (R Development CoreTeam; http://www.r-project.org).

Follow-up
A clinical evaluation was performed after radiotherapy every 3 months the first 2 years then every 6 months. Database was locked on May 30, 2016.

RESULTS
Clinical Outcome
The median follow-up was 38 months (range, 2–80 months). The 2-year DFS was 56.4% (95% confidence interval [CI], 47.3%–67.3%), and the 2-year LRC was 60.7% (95% CI, 51.6%–71.3%). At the analysis, 44 patients had died, and 47 presented a recurrence (20 with locoregional recurrence, 13 with distant recurrence, and 14 with both locoregional and distant recurrence).

Predictive Parameters
Figure 2 shows the correlation between the volume of MTV and the chosen threshold. No limitation in computation of MTV due to the use of an ROI was found.

\(^{23}\)SUVmax was not correlated with LRC or DFS (C-index = 0.54, \( P = 0.63 \)). No difference was found between MTV and TLG. All thresholds between 40% and 60% of SUVmax or between 4.5 and 6 mg/mL appear to have a similar predictive value (Fig. 3). Relative thresholds lower than 36% or higher than 84% were not significantly correlated with DFS (Fig. 4). The best threshold to predict OS and DFS was 51% of SUVmax. (C-index = 0.68 for OS [hazard ratio, 1.43 per 10 mL; 1.23–1.65; \( P < 0.001 \); and C-index = 0.65 for DFS [hazard ratio, 1.43 per 10 mL; 1.23–1.65; \( P = 0.03 \)].) Gross tumor volume was also correlated with DFS (C-index = 0.66, \( P = 0.04 \)) and LRC (C-index = 0.66, \( P = 0.03 \)). In multivariate analysis, MTV 51% was the only significant parameter.

The estimated cutoff point by the Hothorn and Lausen\(^{24}\) method for the MTV 51% was 22.7 mL. Based on this cutoff, 2 risk groups were identified. The 2-year DFS and LRC were 63.3% (95% CI, 53.2%–75.5%) and 68% (95% CI, 48%–79.7%) for the group with MTV 51% of less than 22.7 mL versus 32.9% (95% CI, 18.7%–58.1%) (\( P < 0.001 \)) and 35.3% (95% CI, 20.4%–61.2%) (\( P < 0.001 \)) for the group with MTV 51% of 22.7 mL or greater (\( P = 0.004 \)) (Fig. 5), respectively.

DISCUSSION
To the best of our knowledge, our study is the first one addressing the issue of the predictive value of a wide range of different thresholds (from 0% to 100% of SUVmax) of MTV and TLG in the specific context of oropharyngeal cancers. Considering both primary tumor and lymph node, we found that a relative threshold of 51% was the best predictor for OS and DFS. However, all thresholds between 40% and 62% of SUVmax or between 4.5 and 6 mg/mL appear to have a similar predictive value. The most predictive threshold was 51%, whereas GTV, SUVmax, and parameters computed from absolute SUV threshold appear less predictive. The use of a relative threshold rather than an absolute
This PET parameter may be noteworthy, the reproducibility of the MTV and/or TLG is limited by the initial definition of these parameters, which is based on a threshold of SUV; absolute (all pixels with SUV value > x) or relative (all pixels with SUV value > xx % of SUVmax), and most of the studies used only 1 threshold (2.5 or 3 g/mL, or 40%–50%). Six studies compared only 3 or 4 different thresholds of MTV and/or TLG, most often using the same thresholds of 40%, 50%, and 2.5 and 3 of absolute SUV$^4$–$^{12,13,27,28}$. An absolute threshold of 2.5$^{10,17}$ and a relative threshold of 40%$^{9,12}$ were the best predictors for OS and DFS. However, also, all the other studied thresholds were correlated with OS and DFS but with a lower predictive value. Our study confirms that the use of different relative thresholds within a reasonable range (between 40% and 60%) seems to have no major impact on the predictive value of MTV.

Regarding absolute thresholds, we found a higher value (from 4.5 to 6) than the 2.5 value used routinely. However, same study in cervical cancer. This monocentric study compared 14 thresholds (from 2.0 to 7, and 30%, 40%, and 50% of SUVmax) in 80 patients with head and neck cancer treated with surgery and/or radiotherapy. An absolute threshold of 5 was the best one to predict recurrence and death in head and neck cancer. However, the authors computed MTV only for the primary tumor, and not for the lymph nodes.

Another controversial issue is the use of a cutoff value for MTV, which largely varied from 4.9 to 65 mL (median, 13.1 mL). The use of different thresholds made it difficult to identify the best cutoff to predict clinical outcome. In Abgral et al., a cutoff of 4.9 mL for the MTV 5 was used. However, univariate and multivariate analyses were performed using dichotomized parameters instead of continuous parameters. Dichotomization leads to loss of power, affects the ability to detect relationships, and overestimates the effect size. In our study, in a first step, we used continuous parameters to identify the best threshold (MTV 51%), and in a second step, we used the Hothorn and Lauen method$^{24}$ to determine the best cutoff (23 mL).

Our study had some limitations. It was a retrospective analysis; therefore, without independent validation. We also calculated MTV with the same threshold for both primary tumor and lymph nodes, instead of using a combination of different thresholds, which may have provided a better predictive value. Despite these limitations, we showed that MTV is an independent prognostic factor, with a higher predictive value than SUVmax and GTV.

**CONCLUSIONS**

The use of different thresholds within a reasonable range (between 5.5 and 7 for an absolute threshold and between 40% and 68% for a relative threshold) seems to have no major impact on the predictive value of PET parameters. Metabolic tumor volume for both primary tumor and lymph node computed with a relative threshold of 51% of SUV max was the best predictor of OS and DFS. This parameter may be used to identify patients with a high risk of recurrence of death and who may benefit from treatment intensification.

**REFERENCES**


