Title: Overview of the predictive value of quantitative 18 FDG PET in head and neck cancer treated with chemoradiotherapy.

Authors: Castelli J, De Bari B, Depeursinge A, Simon A, Devillers A, Roman Jimenez G, Prior J, Ozsahin M, de Crevoisier R, Bourhis J

Journal: Critical reviews in oncology/hematology

Year: 2016 Dec

Issue: 108

Pages: 40-51

DOI: 10.1016/j.critrevonc.2016.10.009

Overview of the predictive value of quantitative 18 FDG PET in head and neck cancer treated with chemoradiotherapy

J.Castelli[1,2,3]; B. De Bari[1]; A.Depeursinge[4,5]; A.Simon[2,3]; A.Devillers[6]; G.Roman Jimenez[2,3,7] ; J.Prior[8] ; M.**Ozsahin** [1]; R.de Crevoisier[2,3,9]; J.Bourhis[1]

1. Radiotherapy Department, CHUV, Lausanne, Suisse
2. INSERM, U1099, Rennes, F-35000, France
3. Université de Rennes 1, LTSI, Rennes, F-35000, France
4. University of Applied Sciences Western Switzerland, 3960 Sierre, Switzerland
5. Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne VD, Switzerland
6. Nuclear Medecine Department, Centre Eugene Marquis, Rennes, F-35000, France
7. Keosys medical imaging, 1 impasse Augustin Fresnel, Saint-herblain, F-44815, France
8. Nuclear Medecine Department, CHUV, Suisse
9. Radiotherapy Department, Centre Eugene Marquis, Rennes, F-35000, France

Corresponding author: Dr Joël Castelli, Department of Radiation Oncology, Centre Eugene Marquis, avenue de la Bataille Flandre Dunkerque, F-35000 Rennes, France;

Tel: +33 29 925 30 20, Fax: +33 29 925 30 05 E-mail address: j.castelli@rennes.unicancer.fr

Key points

- SUVMax is a poor predictive parameter when compared with volumetric parameters (MTV, TLG)
- MTV and/or TLG pre-treatment are well correlated with clinical outcome
- MTV had a better predictive value than GTV and/or AJCC staging

Abstract

18 F-fluorodeoxyglucose (18F-FDG) positron emission tomography/computed tomography (PET/CT) allows to quantify the metabolic activity of a tumor (glycolysis) and has become a reference tool in oncology for the staging, restaging, radiotherapy planning and monitoring response in many cancers. Quantitative analyses have been introduced in order to overcome some of the limits of the visual methods, allowing an easier and more objective comparison of the inter- and intra-patients variations. The aims of this review were to report available evidences on the clinical value of quantitative PET/CT parameters in HNC.

Forty-five studies, for a total of 2928 patients, were analyzed. Most of the data available dealt with the intensity of the metabolism, calculated from the Standard Uptake Value (SUV). Metabolic Tumor Volume (MTV) was well correlated with overall survival and disease free survival, with a higher predictive value than the maximum SUV. Spatial distribution of metabolism and textural analyses seems promising.

Key words: PET, Head and Neck cancer, chemoradiotherapy, clinical outcome

1 Introduction

Head and neck cancers are among the most common in the world (5th leading cancer by incidence (Parkin et al., 2005)). The American Joint Committee on Cancer (AJCC) staging is generally used to estimate the prognosis and guide therapy (Edge and Compton, 2010). Radio-chemotherapy is a standard treatment of unresectable and/or locally advanced Head and Neck Cancers (Pignon et al., 2000; St Guily et al., 2010). Despite this treatment, the prognosis remains worst and loco-regional recurrence may occur in up to 40% patients, mostly within the first 2-years after treatment (Chajon et al., 2013).^{18}F-fluorodeoxyglucose (^{18}F-FDG) positron emission tomography/computed tomography (PET/CT) allows to quantify the metabolic activity of a tumor (glycolysis) and has become a reference tool in oncology for the staging, radiotherapy planning and monitoring tumor response in many cancers (Cacicedo et al., 2016; Fletcher et al., 2008). For primary tumor diagnosis, ^{18}F-FDG-PET imaging showed a significant better sensitivity (93% vs 65%) and specificity (70% vs 56%) over CT (Gambhir et al., 2001). PET imaging allows a more accurate nodal staging of locally advanced head and neck cancer (Kyzas et al., 2008; Yoo et al., 2013), and could result in changing the therapeutic management in nearly 15% of patients (Lonneux et al., 2010). For patients with cervical node metastases of unknown primary, PET/CT detected a primary tumor in nearly 30% of patients (Rudmik et al., 2011; Wong et al., 2012; Zhu and Wang, 2013).

Thanks to these potential advantages, PET/CT is recommended for the initial staging and for the treatment decision algorithm of advanced head and neck cancer (Yoo et al., 2013). However, in almost all of these studies, only a visual analysis of PET/CT by physician, based on contrast in uptake between normal tissues and potential tumor (i.e. operator dependent), was performed. Visual analysis was sufficient for diagnosis, staging and detection of recurrence, but with the goal of predicting patient' outcome, quantification is necessary. More recently, quantitative analyses have been introduced in order to overcome some of the limits of the

visual methods (Table 1). Indeed, quantitative analysis is less operator dependent than visual analysis and can be fully automated, allowing an easier and more objective comparison of the inter- and intra-patient variations. The main goal of the quantification is to obtain parameters reflecting the tumor activity and/or having a prognostic value.

The aims of this review were to report available evidences on the value of quantitative parameters from PET/CT performed at the diagnosis, during treatment and during follow to predict overall- and disease free survival in head and neck cancer and to discuss theirs limits.

2 Materials and methods

We performed a systematic electronic search of articles published in PubMed/MEDLINE from January 2000 to march 2016. Our search was restricted to articles reporting data obtained on humans and to English-written articles dealing with locally advanced head and neck cancer and PET/CT. All the articles which did not report data on the prognostic value of PET/CT-related parameters were excluded as well as all the articles which reported data obtained only from visual analyses. Hence this review was focused on the prognostic value of parameters obtained from quantitative or semi-quantitative analyses. We included all the studies reporting data on PET/CT performed before, during or after exclusive RT +/- CT, excluding those reporting data from surgical series and/or post-operative radio-chemotherapy. The predictive value of PET at diagnosis, during treatment and during follow up was analyzed separately.

3 Results

One hundred and twenty-five studies were identified according to the criteria described above. Seventy-seven studies were excluded since they did not match the inclusion criteria, mainly because they dealt with operated patients (22/77 studies). One retrospective study

presenting data on a small population (< 20 patients) was also excluded. Finally, 45 studies were included in the analysis, for a total of 2928 patients. Table 2 summarizes the main characteristics of the studies included in this analysis, while table 3 summarizes the principal results of these studies.

3.1 Predictive value of 18FDG PET before treatment with RT-CT

Forty-two studies investigated the predictive value of quantitative PET parameters at diagnosis (Table 4). The large majority of these studies analyzed parameters based on Standard Uptake Value (SUV), while only 3 studies performed texture or shape analysis.

3.1.1 SUVmax and Metabolic Tumor Volume

Maximum standard uptake value (SUVMax) corresponding to the maximal pixel value in the tumor. Thanks to its ease of use, it was historically the first parameter analyzed. SUVmax was correlated with overall- or disease free survival in 11 studies (Allal et al., 2002; Brun et al., 2002; Castaldi et al., 2012; Chen et al., 2014; Farrag et al., 2010; Higgins et al., 2012; Kitagawa et al., 2003; Machtay et al., 2009; Matoba et al., 2015; Rasmussen et al., 2015; Sanghera et al., 2005). SUVmax allows to identify patients with a high risk of events (death or recurrence). For example, (Rasmussen et al., 2015) analyzed 287 patients with locally advanced head and neck cancers treated with radiotherapy +/- chemotherapy. SUVmax showed a higher predictive value for recurrence than T stage, N stage and age. The authors developed a prognostic model of freedom from failure at 2 years, in which including SUVmax significantly increased the predictive value, changing the estimated risk by more than 10% for 23% of the patients. In (Allal et al., 2002), 63 patients treated with RT +/- CT were prospectively included. Patients presenting a SUVMax < 5.5 g/ml had a 3-year DFS of 79% compared to 42% for those

with SUVmax > 5.5 g/ml (p=0.005). However, the range of cutoff values adopted in published studies to define patients at high or low risk of events markedly varied between 3.7 and 9 g/ml (median: 5.8). Noteworthy, also 2 negative studies are available in the literature (Ashamalla et al., 2014; Greven et al., 2001). In (Greven et al., 2001), patients with local recurrence had a mean pretreatment SUV_{Max} of 7.7 g/ml versus 8.2 g/ml for patients without local recurrence. In (Ashamalla et al., 2014), SUVMax was correlated with OS in univariate analysis, but not in multivariate analysis. Only 28 patients were included in this study, which may explain this negative result.

The Metabolic Tumor Volume (MTV), defined as the volume of FDG activity in a tumor assessed by automated volume of interest delineation, and Total Lesion Glycolysis (TLG), defined as MTV x SUVmean, may be more representative of the tumor heterogeneity. The predictive value of MTV was evaluated in 26 studies, with 21 of them also evaluating SUVMax (for a total of 1464 patients). All these studies showed that MTV/TLG were predictive for clinical outcome, with a higher predictive value than SUVmax. In (Chang et al., 2012), 108 patients with nasopharyngeal cancer treated with RT-CT were prospectively included to assess the predictive value of SUVMax, MTV and TLG for DFS and OS. Only Epstein–Barr virus DNA load and TLG of the tumor were significantly correlated with DFS and OS. In particular, patients presenting a TLG value < 65 g showed a 3-year DFS of 79.9% versus 37.4% for other patients (p<0.001), with a hazard ratio of 3.54 (p=0.006) for DFS and of 4.91 (p=0.045) for OS.

In two studies, MTV was found to have a higher predictive value than TNM staging (Kao et al., 2012; Romesser et al., 2014). (Romesser et al., 2014) reported data of 100 oropharyngeal cancer, treated with RT-CT (median follow-up: 49 months). MTV at a cutoff of 9.7 mL was correlated with DFS (80.3% vs 56.7%, p=0.015) and OS (84.1% vs 57.8% p=0.008). In multivariate analysis, only MTV was significant while GTV, T stage and N stage did not.

Noteworthy, the reproducibility of the MTV and/or TLG may be 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). The choice of the threshold for either method may affect the absolute value of the MTV. Six studies compared the predictive value of MTV and/or TLG computed with different thresholds (Cheng et al., 2015; Kao et al., 2012; Lin et al., 2015; Schinagl et al., 2011; Yabuki et al., 2015). In the study by (Schinagl et al., 2011), 4 thresholds (2.5, 40%, 50% and adaptive threshold based on liver uptake) were compared for 77 patients treated with RT +/- CT. MTV 40% was the strongest predictor of DFS and OS. However, even if the predictive value of the other thresholds was slightly lower, they were also correlated with OS and DFS. Same results were reported by the others studies. Based on these results, the use of different thresholds within a reasonable range (between 2 and 3 for an absolute threshold; and between 40 – 50 % for a relative threshold) seems to have no major impact on the predictive value of MTV.

3.1.2 Texture and shape analysis

Two different approaches have been used to evaluate tumor heterogeneity, one morphological at macroscopic level (shape of the metabolic area) and the other at pixel level (texture analysis). (Apostolova et al., 2014) used a new parameter to characterize the deviation of the tumor's shape from sphere symmetry (asphericity). The initial assumption of the authors was that "aggressive" tumors are expected to show more irregular shapes, due to necrosis, angiogenesis and extravascular extracellular matrix. In a first study, including patients treated with surgery, radiotherapy or chemotherapy alone, asphericity was correlated with OS and PFS. Based on these results, the authors tried to confirm the predictive value of asphericity in a following study (Hofheinz et al., 2015). Thirty-three patients, with

LAHNC treated with RT-CT were included. Using the same cutoff of 20.4 found in (Apostolova et al., 2014), asphericity was correlated with PFS (HR 2.96, p= 0.015) and OS (HR 5.9, p=0.001).

Two studies evaluated the prognostic value of texture analysis in LAHNC. In the first study (Cheng et al., 2013), including 70 oropharyngeal cancers, TLG and texture uniformity were correlated with OS (HR 5.85 and 0.46 respectively). A 3-point risk scale for DFS and OS was proposed, according to the presence of a uniformity ≤ 0.138 and a TLG > 122.9 g. One point was given for each factor. Clinical outcome (DFS or OS) was significantly different in the 3 risk groups. These findings were confirmed in an independent series of 88 oropharyngeal cancer patients (Cheng et al., 2015).

3.2 Predictive value of quantitative PET parameters during chemoradiotherapy

Early changes in tumor metabolism during radiochemotherapy may be assessed by PET/CT and may be used to tailor treatment. The aims of this adaptive strategy to the treatment' response are to decrease the adverse effect and/or to intensify the treatment, with the final goal to improve the outcome.

Seven studies (374 patients) evaluated the predictive value of PET performed during RT +/- CT (Brun et al., 2002; Castaldi et al., 2012; Chen et al., 2014; Farrag et al., 2010; Hentschel et al., 2011; Min et al., 2016; Min et al., 2015). All but one of them found a correlation between PET parameters RT +/- CT and clinical outcome. In a study by Min et al., 100 patients received a PET before and 3 weeks after the beginning of treatment (Min et al., 2016). The authors showed that pre-treatment SUVMax and mid-treatment TLG were correlated with 2-year DFS in multivariate analysis (83% vs 71.4%, p=0.0019 and 88.4% vs 77.2%, p=0.012, respectively). Moreover, patients presenting pretreatment TLG < 91 g and a mid-treatment TLG < 9.4 g presented a better 2-year DFS (88.1% vs 61.1%, p=0.001) and 2-year OS (90% vs 67%, p=0.012).

Other parameter,such as SUVMax and MTV were also correlated to DFS and OS, but TLG was the most predictive one. In (Castaldi et al., 2012), which included 24 patients, no predictive value of PET during treatment was shown. However, the decrease of SUVmax between PET at diagnosis and during treatment was highly correlated with 2-year DFS (100% in case of complete response vs 74% in case of partial response, defined as a reduction of 25% in tumor 18FDG SUV (Young et al., 1999)).

The optimal time to perform PET during treatment is still unclear. Most of the studies performed the PET before the third week to allow time for adapting therapy. A prospective multicentric study (TEMPORAL) (NCT02469922) is undergoing to assess the predictive value of PET at the 2nd and 4th week of chemoradiotherapy. One hundred twenty-three patients are expected to be included.

3.3 Predictive value of 18FDG PET after treatment

After treatment with radiotherapy, PET/CT may be used to identify good responders and avoid useless neck dissection. Twelve studies performed a quantitative or semi quantitative analysis from PET after treatment. All these studies evaluated the SUVMax. A high SUVMax in post treatment was correlated with a poor outcome in 6 studies (Horiuchi et al., 2008; Hoshikawa et al., 2011; Ito et al., 2014; Kim et al., 2016; Kitagawa et al., 2003; Moeller et al., 2010). In (Moeller et al., 2010), 98 patients underwent a PET before and 8 weeks after RT +/- CT. The authors found that a post-treatment SUVMax $\leq$ 6 g/ml and the variation of SUVMax (in %) between the pre- and post-irradiation PET/CT were predictive for DFS. In (Kim et al., 2016), a PET was performed 3 months after RT-CT. Seventy-eight patients were analyzed. Three-year OS was 87.7% in patients with SUVMax <4.4 g/ml versus 56.9% (p=0.002).

A comparison between visual analysis and quantitative parameters was performed by (Hoshikawa et al., 2009). Thirty-five patients underwent PET before and 5 weeks after RT-CT. Patients with a post-treatment SUVMax value > 3 g/ml and decreasing less than 60% compared to the pre-treatment situation presented a higher risk of recurrence (odds ratio = 61.5, $p<0.0001$). The overall accuracy for quantitative analysis was 89.9% vs 60.9% for the visual analysis.

4 Discussion and conclusion

This overview of the available literature shows that MTV and TLG are well correlated with clinical outcome (Local control, Disease Free Survival and overall survical). Most of the available data deal with the intensity of the metabolism, calculated from the SUV, a quantitative parameter used to normalize the uptake of ^{18}F-FDG. In practice, SUV is defined as a ratio of tissue radioactivity concentration and the injected dose adjusted by body weight (SUV_{bw} with BW for body weight). Intensity of the metabolism can be analyzed using histogram-based method, which represents the voxel value frequency distribution. This method includes in particular the four histogram moments, i.e., the mean (corresponding to SUVMean), the maximum (corresponding to SUVMax), the median, the skewness (asymmetry of the histogram) and kurtosis (degree of peakedness of a distribution). However, it did not take into account the spatial relationship between voxel values (Fig. 1).

The maximum SUV (SUVMax) corresponds to the maximal pixel value in the tumor. Thanks to its ease of use, it is one of the most used parameters in the clinical practice. However, this value is highly dependent from noise, duration and parameters of acquisition, and so is considered to be poorly reproducible (Boellaard et al., 2004; Nahmias and Wahl, 2008; Nakamoto et al., 2002). This point may explain the wide range of cut-off value for SUVmax

reported in the available studies (from 3.7 to 9 g/ml), limiting the generalization of the use of SUVmax for the whole population. Peak SUV (SUV_{Peak}), defined as the average SUV within a small region of interest (1.2 cm of diameter) around the SUV_{Max}, is a more robust alternative to SUV_{Max}. However, SUV_{Peak} may not be representative of nonhomogeneous overall tumor uptake, and the ideal size of the ROI is still unclear (Lee et al., 2007).

Others volumetric parameters like the MTV, the mean SUV within the tumor volume or the TLG are used to represent the heterogeneity of the tumor uptake. The predictive value for clinical outcome of these parameters seems to be higher than SUV_{Max}. However, uptake in PET may be due to inflammatory or infectious reaction. Furthermore, the physiological uptake surrounding organs can also be a source of loss of specificity in the analysis of the signal. A major difficulty in the analysis of PET is to differentiate the tumor signal from the non-tumor signal. PET imaging suffers from a low contrast and spatial resolution, with a high noise background and partial effect volume. Tumor delineation may change depending on the chosen segmentation method. One of the most used automatic method is to use a threshold, between 2 and 3 (absolute value) or 40-50% (relative value of SUVmax).

One important issue concerning the predictive value of MTV is the lack of external validation. Most of the studies were monocentric, using the same PET/CT for all patients. Only two studies performed a validation on an independent dataset (Hofheinz et al., 2015; Tang et al., 2012). The first study (La et al., 2009) included 85 patients and showed that an increase of MTV of 17 cm^3 (from the 25th to 75th percentile) was significantly correlated with an increased risk of death (HR 2.1). The authors validated their results on a dataset of 83 patients treated in the same institution after the original dataset (Tang et al., 2012). Based on (Apostolova et al., 2014), (Hofheinz et al., 2015) used a cutoff of TLG of 58.7 ml. They showed a correlation of TLG only with better DFS (HR 3.01, p=0.048) but not with better OS (HR 2.02, p=0.22). After adjusting the cutoff at a value of 141 ml, TLG was also correlated with OS (HR

3.32, p=0.016). Such methodologies and findings highlight the difficulty in identifying a cutoff which may be tested on external dataset of patients. The use of international guidelines, like the European Association of Nuclear Medecine guidelines for tumor imaging (Boellaard et al., 2015), by harmonizing quantitative FDG PET/CT imaging procedures in multicentre studies and quantitative interpretation criteria, may increases the reproducibility of PET studies.

The spatial relationships between the voxel values within the tumor may be assessed by texture analyses. Texture analyses aim to characterize the internal metabolism morphology of the tumors. From a technical point of view, they characterize the transitions between voxel values. Several approaches exist and they all rely on a quantification of spatial scales organization and directions in images. Most approaches are computing the latter in two dimensions on a slice basis. The most widely used method is the Gray Level Co-occurrence Matrix and consists in calculating matrices counting the co-occurrences of two voxels values separated from a set of fixed distances and along set of fixed directions. Several statistics can be computed on these matrices to quantify textural properties (e.g., correlation, contrast, energy). Another popular approach is to apply image filters with various scales and directional properties to continuously quantify transitions between image voxels. One popular example is the isotropic Mexican hat filter (also called Laplacian of Gaussian filter) . A comprehensive review of methods for 3-D texture analysis is available in (Depeursinge et al., 2014). This kind of analysis seems promising, but its use should still be considered experimental and limited to clinical studies.

Unresolved questions and controversies

1. Reproducibility of PET parameters between different machines and/or centre
2. Which methods (Manually, relative, absolute or adaptive segmentation) and which threshold to compute MTV ?
3. When should perform PET during radiotherapy ?
4. Which methods for texture and shape analysis ?

Conflicts of interest: none

Acknowledgments

The authors wish to thank Ms S.Colombe who assisted in proofreading the manuscript.

5 References

Allal, A.S., Dulguerov, P., Allaoua, M., Haenggeli, C.A., El-Ghazi el, A., Lehmann, W., Slosman, D.O., 2002. Standardized uptake value of 2-[(18)F] fluoro-2-deoxy-D-glucose in predicting outcome in head and neck carcinomas treated by radiotherapy with or without chemotherapy. J Clin Oncol 20(5), 1398-1404.

Apostolova, I., Steffen, I.G., Wedel, F., Lougovski, A., Marnitz, S., Derlin, T., Amthauer, H., Buchert, R., Hofheinz, F., Brenner, W., 2014. Asphericity of pretherapeutic tumour FDG uptake provides independent prognostic value in head-and-neck cancer. Eur Radiol 24(9), 2077-2087.

Ashamalla, H., Mattes, M., Guirguis, A., Zaidi, A., Mokhtar, B., Tejwani, A., 2014. The Anatomical Biological Value on Pretreatment (18)F-fluorodeoxyglucose Positron Emission Tomography Computed Tomography Predicts Response and Survival in Locally Advanced Head and Neck Cancer. World J Nucl Med 13(2), 102-107.

Boellaard, R., Delgado-Bolton, R., Oyen, W.J., Giammarile, F., Tatsch, K., Eschner, W., Verzijlbergen, F.J., Barrington, S.F., Pike, L.C., Weber, W.A., Stroobants, S., Delbeke, D., Donohoe, K.J., Holbrook, S., Graham, M.M., Testanera, G., Hoekstra, O.S., Zijlstra, J., Visser, E., Hoekstra, C.J., Pruim, J., Willemsen, A., Arends, B., Kotzerke, J., Bockisch, A., Beyer, T., Chiti, A., Krause, B.J., European Association of Nuclear, M., 2015. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging 42(2), 328-354.

Boellaard, R., Krak, N.C., Hoekstra, O.S., Lammertsma, A.A., 2004. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. J Nucl Med 45(9), 1519-1527.

Brun, E., Kjellen, E., Tennvall, J., Ohlsson, T., Sandell, A., Perfekt, R., Perfekt, R., Wennerberg, J., Strand, S.E., 2002. FDG PET studies during treatment: prediction of therapy outcome in head and neck squamous cell carcinoma. Head & neck 24(2), 127-135.

Cacicedo, J., Navarro, A., Del Hoyo, O., Gomez-Iturriaga, A., Alongi, F., Medina, J.A., Elicin, O., Skanjeti, A., Giammarile, F., Bilbao, P., Casquero, F., de Bari, B., Dal Pra, A., 2016. Role of fluorine-18 fluorodeoxyglucose PET/CT in head and neck oncology: the point of view of the radiation oncologist. Br J Radiol, 20160217.

Castaldi, P., Rufini, V., Bussu, F., Micciche, F., Dinapoli, N., Autorino, R., Lago, M., De Corso, E., Almadori, G., Galli, J., Paludetti, G., Giordano, A., Valentini, V., 2012. Can "early" and "late"18F-FDG PET-CT be used as prognostic factors for the clinical outcome of patients with locally advanced head and neck cancer treated with radio-chemotherapy? Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 103(1), 63-68.

Chajon, E., Lafond, C., Louvel, G., Castelli, J., Williaume, D., Henry, O., Jegoux, F., Vauleon, E., Manens, J.P., Le Prise, E., de Crevoisier, R., 2013. Salivary gland-sparing other than parotid-sparing in definitive head-and-neck intensity-modulated radiotherapy does not seem to jeopardize local control. Radiat Oncol 8, 132.

Chang, K.P., Tsang, N.M., Liao, C.T., Hsu, C.L., Chung, M.J., Lo, C.W., Chan, S.C., Ng, S.H., Wang, H.M., Yen, T.C., 2012. Prognostic significance of 18F-FDG PET parameters and plasma Epstein-Barr virus DNA load in patients with nasopharyngeal carcinoma. J Nucl Med 53(1), 21-28.

Chen, S.W., Hsieh, T.C., Yen, K.Y., Yang, S.N., Wang, Y.C., Chien, C.R., Liang, J.A., Kao, C.H., 2014. Interim FDG PET/CT for predicting the outcome in patients with head and neck cancer. The Laryngoscope 124(12), 2732-2738.
Cheng, N.M., Fang, Y.H., Chang, J.T., Huang, C.G., Tsan, D.L., Ng, S.H., Wang, H.M., Lin, C.Y., Liao, C.T., Yen, T.C., 2013. Textural features of pretreatment 18F-FDG PET/CT images: prognostic significance in patients with advanced T-stage oropharyngeal squamous cell carcinoma. J Nucl Med 54(10), 1703-1709.
Cheng, N.M., Fang, Y.H., Lee, L.Y., Chang, J.T., Tsan, D.L., Ng, S.H., Wang, H.M., Liao, C.T., Yang, L.Y., Hsu, C.H., Yen, T.C., 2015. Zone-size nonuniformity of 18F-FDG PET regional textural features predicts survival in patients with oropharyngeal cancer. Eur J Nucl Med Mol Imaging 42(3), 419-428.
Depeursinge, A., Foncubierta-Rodriguez, A., Van De Ville, D., Muller, H., 2014. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. Med Image Anal 18(1), 176-196.
Edge, S.B., Compton, C.C., 2010. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 17(6), 1471-1474.
Farrag, A., Ceulemans, G., Voordeckers, M., Everaert, H., Storme, G., 2010. Can 18F-FDG-PET response during radiotherapy be used as a predictive factor for the outcome of head and neck cancer patients? Nuclear medicine communications 31(6), 495-501.
Fletcher, J.W., Djulbegovic, B., Soares, H.P., Siegel, B.A., Lowe, V.J., Lyman, G.H., Coleman, R.E., Wahl, R., Paschold, J.C., Avril, N., Einhorn, L.H., Suh, W.W., Samson, D., Delbeke, D., Gorman, M., Shields, A.F., 2008. Recommendations on the use of 18F-FDG PET in oncology. J Nucl Med 49(3), 480-508.
Gambhir, S.S., Czernin, J., Schwimmer, J., Silverman, D.H., Coleman, R.E., Phelps, M.E., 2001. A tabulated summary of the FDG PET literature. J Nucl Med 42(5 Suppl), 1S-93S.
Ganeshan, B., Miles, K.A., Young, R.C., Chatwin, C.R., 2008. Three-dimensional selective-scale texture analysis of computed tomography pulmonary angiograms. Invest Radiol 43(6), 382-394.
Greven, K.M., Williams, D.W., 3rd, McGuirt, W.F., Sr., Harkness, B.A., D'Agostino, R.B., Jr., Keyes, J.W., Jr., Watson, N.E., Jr., 2001. Serial positron emission tomography scans following radiation therapy of patients with head and neck cancer. Head & neck 23(11), 942-946.
Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics SMC-3(6), 610-621.
Hentschel, M., Appold, S., Schreiber, A., Abolmaali, N., Abramyuk, A., Dorr, W., Kotzerke, J., Baumann, M., Zophel, K., 2011. 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 38(7), 1203-1211.
Higgins, K.A., Hoang, J.K., Roach, M.C., Chino, J., Yoo, D.S., Turkington, T.G., Brizel, D.M., 2012. Analysis of pretreatment FDG-PET SUV parameters in head-and-neck cancer: tumor SUVmean has superior prognostic value. Int J Radiat Oncol Biol Phys 82(2), 548-553.
Hofheinz, F., Lougovski, A., Zophel, K., Hentschel, M., Steffen, I.G., Apostolova, I., Wedel, F., Buchert, R., Baumann, M., Brenner, W., Kotzerke, J., van den Hoff, J., 2015. Increased evidence for the prognostic value of primary tumor asphericity in pretherapeutic FDG PET for risk stratification in patients with head and neck cancer. Eur J Nucl Med Mol Imaging 42(3), 429-437.
Horiuchi, C., Taguchi, T., Yoshida, T., Nishimura, G., Kawakami, M., Tanigaki, Y., Matsuda, H., Mikami, Y., Oka, T., Inoue, T., Tsukuda, M., 2008. Early assessment of clinical response to concurrent chemoradiotherapy in head and neck carcinoma using fluoro-2-deoxy-d-glucose positron emission tomography. Auris Nasus Larynx 35(1), 103-108.
Hoshikawa, H., Kishino, T., Nishiyama, Y., Yamamoto, Y., Yonezaki, M., Mori, N., 2011. Early prediction of local control in head and neck cancer after chemoradiotherapy by FDG-PET. Nuclear medicine communications 32(8), 684-689.
Hoshikawa, H., Mitani, T., Nishiyama, Y., Yamamoto, Y., Ohkawa, M., Mori, N., 2009. Evaluation of the therapeutic effects and recurrence for head and neck cancer after chemoradiotherapy by FDG-PET. Auris Nasus Larynx 36(2), 192-198.
Ito, K., Shimoji, K., Miyata, Y., Kamiya, K., Minamimoto, R., Kubota, K., Okasaki, M., Morooka, M., Yokoyama, J., 2014. Prognostic value of post-treatment (18)F-FDG PET/CT for advanced head and neck cancer after combined intra-arterial chemotherapy and radiotherapy. Chin J Cancer Res 26(1), 30-37.

Kao, C.H., Lin, S.C., Hsieh, T.C., Yen, K.Y., Yang, S.N., Wang, Y.C., Liang, J.A., Hua, C.H., Chen, S.W., 2012. Use of pretreatment metabolic tumour volumes to predict the outcome of pharyngeal cancer treated by definitive radiotherapy. Eur J Nucl Med Mol Imaging 39(8), 1297-1305.
Keyes, J.W., Jr., 1995. SUV: standard uptake or silly useless value? J Nucl Med 36(10), 1836-1839.
Kim, R., Ock, C.Y., Keam, B., Kim, T.M., Kim, J.H., Paeng, J.C., Kwon, S.K., Hah, J.H., Kwon, T.K., Kim, D.W., Wu, H.G., Sung, M.W., Heo, D.S., 2016. Predictive and prognostic value of PET/CT imaging post-chemoradiotherapy and clinical decision-making consequences in locally advanced head & neck squamous cell carcinoma: a retrospective study. BMC Cancer 16(1), 116.
Kitagawa, Y., Sano, K., Nishizawa, S., Nakamura, M., Ogasawara, T., Sadato, N., Yonekura, Y., 2003. FDG-PET for prediction of tumour aggressiveness and response to intra-arterial chemotherapy and radiotherapy in head and neck cancer. Eur J Nucl Med Mol Imaging 30(1), 63-71.
Kyzas, P.A., Evangelou, E., Denaxa-Kyza, D., Ioannidis, J.P., 2008. 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: a meta-analysis. J Natl Cancer Inst 100(10), 712-720.
La, T.H., Filion, E.J., Turnbull, B.B., Chu, J.N., Lee, P., Nguyen, K., Maxim, P., Quon, A., Graves, E.E., Loo, B.W., Jr., Le, Q.T., 2009. Metabolic tumor volume predicts for recurrence and death in head-and-neck cancer. Int J Radiat Oncol Biol Phys 74(5), 1335-1341.
Lee, P., Weerasuriya, D.K., Lavori, P.W., Quon, A., Hara, W., Maxim, P.G., Le, Q.T., Wakelee, H.A., Donington, J.S., Graves, E.E., Loo, B.W., Jr., 2007. Metabolic tumor burden predicts for disease progression and death in lung cancer. Int J Radiat Oncol Biol Phys 69(2), 328-333.
Lin, Y.C., Chen, S.W., Hsieh, T.C., Yen, K.Y., Yang, S.N., Wang, Y.C., Kao, C.H., 2015. Risk stratification of metastatic neck nodes by CT and PET in patients with head and neck cancer receiving definitive radiotherapy. J Nucl Med 56(2), 183-189.
Lonneux, M., Hamoir, M., Reychler, H., Maingon, P., Duvillard, C., Calais, G., Bridji, B., Digue, L., Toubeau, M., Gregoire, V., 2010. Positron emission tomography with [18F]fluorodeoxyglucose improves staging and patient management in patients with head and neck squamous cell carcinoma: a multicenter prospective study. J Clin Oncol 28(7), 1190-1195.
Machtay, M., Natwa, M., Andrel, J., Hyslop, T., Anne, P.R., Lavarino, J., Intenzo, C.M., Keane, W., 2009. Pretreatment FDG-PET standardized uptake value as a prognostic factor for outcome in head and neck cancer. Head & neck 31(2), 195-201.
Matoba, M., Tuji, H., Shimode, Y., Kondo, T., Oota, K., Tonami, H., 2015. Lesion regression rate based on RECIST: prediction of treatment outcome in patients with head and neck cancer treated with chemoradiotherapy compared with FDG PET-CT. J Radiat Res 56(3), 553-560.
Min, M., Lin, P., Lee, M., Shon, I.H., Lin, M., Forstner, D., Tieu, M.T., Chicco, A., Bray, V., Fowler, A., 2016. 18F-FDG PET-CT performed before and during radiation therapy of head and neck squamous cell carcinoma: Are they independent or complementary to each other? Journal of medical imaging and radiation oncology.
Min, M., Lin, P., Lee, M.T., Shon, I.H., Lin, M., Forstner, D., Bray, V., Chicco, A., Tieu, M.T., Fowler, A., 2015. 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 42(13), 1984-1994.
Moeller, B.J., Rana, V., Cannon, B.A., Williams, M.D., Sturgis, E.M., Ginsberg, L.E., Macapinlac, H.A., Lee, J.J., Ang, K.K., Chao, K.S., Chronowski, G.M., Frank, S.J., Morrison, W.H., Rosenthal, D.I., Weber, R.S., Garden, A.S., Lippman, S.M., Schwartz, D.L., 2010. Prospective imaging assessment of mortality risk after head-and-neck radiotherapy. Int J Radiat Oncol Biol Phys 78(3), 667-674.
Murphy, J.D., La, T.H., Chu, K., Quon, A., Fischbein, N.J., Maxim, P.G., Graves, E.E., Loo, B.W., Jr., Le, Q.T., 2011. Postradiation metabolic tumor volume predicts outcome in head-and-neck cancer. Int J Radiat Oncol Biol Phys 80(2), 514-521.
Nahmias, C., Wahl, L.M., 2008. Reproducibility of standardized uptake value measurements determined by 18F-FDG PET in malignant tumors. J Nucl Med 49(11), 1804-1808.
Nakamoto, Y., Zasadny, K.R., Minn, H., Wahl, R.L., 2002. Reproducibility of common semi-quantitative parameters for evaluating lung cancer glucose metabolism with positron emission tomography using 2-deoxy-2-[18F]fluoro-D-glucose. Mol Imaging Biol 4(2), 171-178.
Parkin, D.M., Bray, F., Ferlay, J., Pisani, P., 2005. Global cancer statistics, 2002. CA Cancer J Clin 55(2), 74-108.

Pignon, J.P., Bourhis, J., Domenge, C., Designe, L., 2000. 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 355(9208), 949-955.
Rasmussen, J.H., Vogelius, I.R., Fischer, B.M., Friborg, J., Aznar, M.C., Persson, G.F., Hakansson, K., Kristensen, C.A., Bentzen, S.M., Specht, L., 2015. Prognostic value of 18F-fludeoxyglucose uptake in 287 patients with head and neck squamous cell carcinoma. Head & neck 37(9), 1274-1281.
Romesser, P.B., Lim, R., Spratt, D.E., Setton, J., Riaz, N., Lok, B., Rao, S., Sherman, E.J., Schoder, H., Lee, N.Y., 2014. The relative prognostic utility of standardized uptake value, gross tumor volume, and metabolic tumor volume in oropharyngeal cancer patients treated with platinum based concurrent chemoradiation with a pre-treatment [(18)F] fluorodeoxyglucose positron emission tomography scan. Oral Oncol 50(9), 802-808.
Rudmik, L., Lau, H.Y., Matthews, T.W., Bosch, J.D., Kloiber, R., Molnar, C.P., Dort, J.C., 2011. Clinical utility of PET/CT in the evaluation of head and neck squamous cell carcinoma with an unknown primary: a prospective clinical trial. Head Neck 33(7), 935-940.
Sanghera, B., Wong, W.L., Lodge, M.A., Hain, S., Stott, D., Lowe, J., Lemon, C., Goodchild, K., Saunders, M., 2005. Potential novel application of dual time point SUV measurements as a predictor of survival in head and neck cancer. Nuclear medicine communications 26(10), 861-867.
Schinagl, D.A., Span, P.N., Oyen, W.J., Kaanders, J.H., 2011. Can FDG PET predict radiation treatment outcome in head and neck cancer? Results of a prospective study. Eur J Nucl Med Mol Imaging 38(8), 1449-1458.
Schwartz, D.L., Harris, J., Yao, M., Rosenthal, D.I., Opanowski, A., Levering, A., Ang, K.K., Trotti, A.M., Garden, A.S., Jones, C.U., Harari, P., Foote, R., Holland, J., Zhang, Q., Le, Q.T., 2015. 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 91(4), 721-729.
St Guily, J.L., Borget, I., Vainchtock, A., Remy, V., Takizawa, C., 2010. Head and neck cancers in France: an analysis of the hospital medical information system (PMSI) database. Head & neck oncology 2, 22.
Tang, C., Murphy, J.D., Khong, B., La, T.H., Kong, C., Fischbein, N.J., Colevas, A.D., Iagaru, A.H., Graves, E.E., Loo, B.W., Jr., Le, Q.T., 2012. Validation that metabolic tumor volume predicts outcome in head-and-neck cancer. Int J Radiat Oncol Biol Phys 83(5), 1514-1520.
Thie, J.A., 2004. Understanding the standardized uptake value, its methods, and implications for usage. J Nucl Med 45(9), 1431-1434.
Wahl, R.L., Jacene, H., Kasamon, Y., Lodge, M.A., 2009. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. J Nucl Med 50 Suppl 1, 122S-150S.
Wong, W.L., Sonoda, L.I., Gharpurhy, A., Gollub, F., Wellsted, D., Goodchild, K., Lemon, C., Farrell, R., Saunders, M., 2012. 18F-fluorodeoxyglucose positron emission tomography/computed tomography in the assessment of occult primary head and neck cancers--an audit and review of published studies. Clin Oncol (R Coll Radiol) 24(3), 190-195.
Yabuki, K., Shiono, O., Komatsu, M., Sano, D., Nishimura, G., Takahashi, M., Taguchi, T., Inoue, T., Oridate, N., 2015. Predictive and prognostic value of metabolic tumor volume (MTV) in patients with laryngeal carcinoma treated by radiotherapy (RT) / concurrent chemoradiotherapy (CCRT). PLoS One 10(2), e0117924.
Yoo, J., Henderson, S., Walker-Dilks, C., 2013. Evidence-based guideline recommendations on the use of positron emission tomography imaging in head and neck cancer. Clin Oncol (R Coll Radiol) 25(4), e33-66.
Young, H., Baum, R., Cremerius, U., Herholz, K., Hoekstra, O., Lammertsma, A.A., Pruim, J., Price, P., 1999. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Journal of Cancer 35(13), 1773-1782.
Zhu, L., Wang, N., 2013. 18F-fluorodeoxyglucose positron emission tomography-computed tomography as a diagnostic tool in patients with cervical nodal metastases of unknown primary site: a meta-analysis. Surgical oncology 22(3), 190-194.

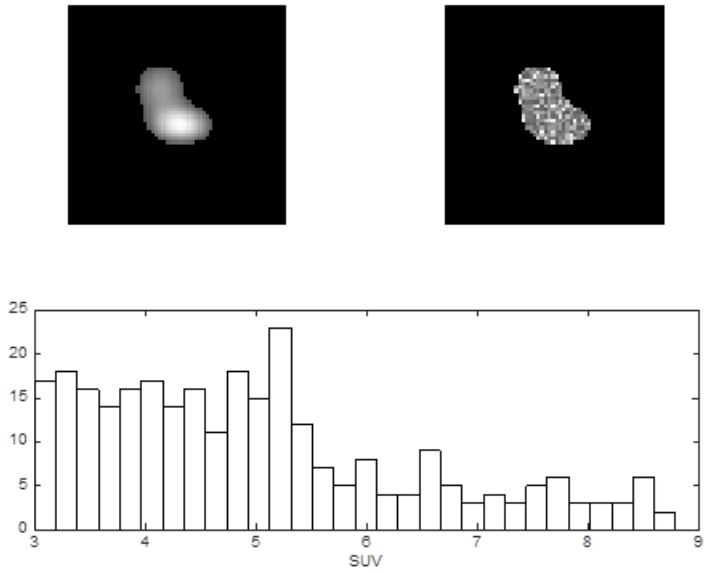

Figure 1 : Heterogeneity measures do not characterize the spatial relationships between voxels. The two tumors in the upper row have identical SUV histograms, although their visual aspect is very dissimilar.

Table 1 : Most frequently used quantitative parameters in PET imaging

Parameters	**Definition and method to compute**
SUV_{Max}	Maximal pixel value in the tumor
SUV Peak	Average SUV within a small, fixed-size region of interest (ROIpeak) of 1.2 cm diameter, centered on a high-uptake part of the tumor
Metabolic Tumor Volume (MTV)	Sum of the volume of voxels with SUV exceeding a certain threshold value in a tumor
SUV Mean	Average SUV in the ROI (defined by applying a threshold or by visual assessment)
Total Lesion Glycolysis (TLG)	TLG is obtained by multiplying MTV and the mean SUV of the MTV

Table 2 : Main criteria of the 45 studies. * = Mean follow up, CR: Complete Response, LCR: Loco Regional Response, LR: Local Relapse, DSS: Disease Specific Survival, RFS: Recurrence Free Survival, LRFS: Local Relapse Free Survival, DFS: Disease Free Survival, OS: Overall Survival, MTV: Metabolic Tumor Volume, TLG: Total Lesion Glycolysis

Authors	Year	Subject No.	Study Design	Timing of te PET-CT	Follow up	Localisation	Treatment	End point	Quantitative PET parameters
Greven [36]	2001	45	Prospective	Pre post	N/A	HNC	RT	LR	SUV_{Max}
Allal [14]	2002	63	Prospective	Pre	36	HNC	RT +/- CT	DFS OS	SUV_{Max}
Brun [19]	2002	47	Prospective	Pre and per	39.6	HNC	RT +/- CT	CR LRC OS	SUV_{Max} Metabolic rate FDG
Kitagawa [16]	2003	20	Prospective	Pre, post	52.8	HNC	CRT	CR	SUV_{Max}
Sanghera [20]	2005	12	Prospective	Pre	24	HNC	RT	OS	SUV_{Max} at 1 and 2h, SUV_{Max} Difference
Horiuchi [43]	2008	31	Retrospective	Pre and post	N/A	HNC	CRT	LR	SUV_{Max}
Chung [61]	2009	82	Retrospective	Pre	34.8*	Pharynx	RT +/-CT	LR DFS OS	SUV_{Max} MTV 2.5
La [34]	2009	85	Retrospective	Pre	20.4 *	HNC	RT +/- CT	OS DSF LRC	SUV_{Max} MTV 50%
Machtay [17]	2009	60	Retrospective	Pre	N/A	HNC	RT +/- CT	OS DFS	SUV_{Max}
Suzuki [62]	2009	45	Retrospective	Pre	24*	HNC	RT	OS DFS	SUV_{Max}
Farrag [15]	2010	43	Prospective	Pre and Per	12.7	HNC	CRT	OS DFS LRRFS	SUV_{Max}
Moeller [44]	2010	98	Prospective	Pre and post	24	HNC	RT +/- CT	DFS	SUV_{Max} T Change in SUV_{Max} T
Seol [63]	2010	59	Retrospective	Pre	N/A	HNC	Neo CT + RT	DFS OS	SUV_{Max} SUVMean MTV 2.5
Deron [64]	2011	22	Retrospective	Pre	20	HNC	RT +/- CT	DFS OS	SUV_{Max} MTV 50%
Hentschel [39]	2011	37	Prospective	Pre and per	26	HNC	CRT	DFS OS LRC	SUV_{Max} SUVMean MTV 50%
Hoshikawa [45]	2011	35	Prospective	Pre and post	50	HNC	RT +/- CT	Recurrence	SUV_{Max}
Murphy [49]	2011	47	Retrospective	Post	34	HNC	CRT	DFS OS	SUV_{Max} MTV 2, 2.5, 3, 3.5, 4 TLG

Schinagl [28]	2011	77	Prospective	Pre	46	HNC	CRT	DFS OS	SUV_{Max} SUVMean MTV 40% 50% MTV 2.5 GTV -PET
Castaldi [21]	2012	24	Prospective	Pre per and post	29.2	HNC	CRT	RFS DSS	SUV_{Max} Change in SUVMax (EORTC criteria)
Chang [25]	2012	108	Prospective	Pre	N/A	Nasopharynx	CRT	OS DFS LRFS	SUV_{Max} MTV 2.5 TLG
Chu [65]	2012	51	Retrospective	Pre	17.5	HNC	RT +/- CT	OS DFS	SUV_{Max} MTV50% MTV Velocity
Higgins [22]	2012	88	Retrospective	Pre	15	HNC	RT +/- CT	DFS LRC OS	SUV_{Max} SUVMean TLG (manuelly delineated)
Kao [27]	2012	64	Retrospective	Pre	24	Pharynx	RT +/- CT	DFS PRFS	MTV 2.5 3.0 40% 50%
Romesser [66]	2012	41	Retrospective	Pre	24.2	HNC	RT +/- CT	OS DFS LRFS	SUV_{Max} MTV (Gradient based method)
Tang [32]	2012	83	Retrospective	Pre	20	HNC	CRT	OS DFS	SUV_{Max} MTV 50%
Cheng [38]	2013	70	Retrospective	Pre	>24	Oropharynx	CRT	OS DFS	MTV 2.5 TLG Textural features
Ashamalla [37]	2014	28	Retrospective	Pre and post	36 *	HNC	RT +/- CT	OS	SUVMax SUVMean Anatomical biological value = SUVMax x greatest tumor diameter
Chen [23]	2014	51	Prospective	Pre and per	23	Pharynx	RT +/- CT	OS DFS	SUV_{Max} pre and per SUV reduction ratio
Hanamoto [67]	2014	118	Prospective	Pre	N/A	HNC	CRT	LR	SUV_{Max} SUVMean MTV 2.5 TLG
Ito [46]	2014	36	Retrospective	Post	23.8 *	HNC	CRT	OS LC	SUV_{Max}
Romesser [26]	2014	100	Retrospective	Pre	49	Oropharyngeal	CRT	LRC DFS OS	SUV_{Max} MTV 42%
Sager [68]	2014	74	Retrospective	Pre	23	HNC	CRT	DFS OS	SUV_{Max} MTV 50%
Akagunduz [69]	2015	62	Retrospective	Pre	18	HNC	RT +/- CT	LRFS DFS OS	SUV_{Max} SULMax MTV (adaptive threshold based)
Cheng [29]	2015	88	Retrospective	Pre	32	Oropharynx	CRT	DFS DSS	MTV 50 % 42% 2.5 and adaptive threshold TLG Textural features
Hofheinz [33]	2015	37	Prospective	Pre	27*	HNC	CRT	DFS OS	SUV_{Max} SUVMean MTV (adaptive threshold) TLG Asphericity

Lin [30]	2015	91	Retrospective	Pre	18	Pharynx	CRT	OS DFS	SUV_{Max} Nodal MTV2.5 N MTV40% N MTV50% N TLG40% N TLG50% N
Matoba [24]	2015	33	Prospective	Pre and post	N/A	HNC	CRT	LRC DFS OS	SUV_{Max} EORTC Criteria
Min [41]	2015	72	Retrospective	Pre and per	25	HNC	CRT	LRFS DFS MFFS OS	SUV_{Max} MTV 2.5 TLG Percentage reduction between per and pre treatment PET
Moon [70]	2015	44	Retrospective	Pre	34.7	Nasopharynx	CRT	DFS	SUV_{Max} SUVMean MTV (adaptive threshold) TLG
Rasmussen [18]	2015	287	Retrospective	Pre	32	HNC	RT +/- CT	Time to failure	SUV_{Max} SUVMean SUVPeak
Schwartz [50]	2015	74	Retrospective	Pre and post	50.4	HNC	CRT	LR DFS OS	SUV_{Max} SUVPeak MTV 40 %
Yabuki [31]	2015	118	Retrospective	Pre	36	Larynx	CRT	OS DFS	SUV_{Max} MTV 2, 2.5, 3
Kim [47]	2016	78	Retrospective	Post	52.7	HNC	CRT	DFS OS	SUV_{Max}
Min [40]	2016	100	Retrospective	Pre and per	20	HNC	RT +/- CT	LRFS DFS MFFS OS	SUV_{Max} MTV 2.5 TLG Percentage reduction between per and pre treatment PET

Table 2 : Correlation between PET quantitative parameters and clinical outcome. CR: Complete Response, LCR: Loco Regional Response, LR: Local Relapse, DSS: Disease Specific Survival, RFS: Recurrence Free Survival, LRFS: Local Relapse Free Survival, DFS: Disease Free Survival, OS: Overall Survival, MTV: Metabolic Tumor Volume, TLG: Total Lesion Glycolysis

Timing	Authors	Year	End point	Used PET parameters	Significant Prognostic parameters	Threshold	Clinical outcome	Hazard Ratio
Pre and post	Greven [36]	2001	LR	SUVMax	None			
Pre	Allal [14]	2002	DFS OS	SUVMax	SUVMax (DFS)	5.5	3-year DFS 79% vs 42%	N/A
Pre and per	Brun [19]	2002	Complete response (CR) LRC OS	SUVmax Metabolic rate (MR) FDG	Pre treatment : SUVMax Tumor (CR and LCR) Per treatment : SUVMax Tumor (CR and LCR), MR Tumor and lymph node (CR and LCR) MR FDG per OS	Pre treatment SUVmaxT = 9 Per trt SUVmaxT = 5	SUVmax pre trt CR 96% vs 64% (p=.01) LRC 96 vs 57 (p=.003) MR Tumor per trt CR 96 vs 62% (p=.007) LRC 96 vs 55% (p=0.002) OS 72% vs 35% (p=0.0042) SUVMax per trt LCR 91% vs 62 (p=.031)	
Pre and post	Kitagawa [16]	2003	Clinical response	SUVMax	SUVMax	N/A	N/A	N/A
Pre	Sanghera [20]	2005	OS	SUVMax at 1 and 2h, SUVMax Difference	SUV difference	16%	N/A	N/A
Pre and post	Horiuchi [43]	2008	LR	SUVMax	SUVMax Post trt	3.7	N/A	
Pre	Chung [61]	2009	LR DFS OS	SUV Max MTV 2.5	MTV	40		DFS 3.42 (p =0.04)
Pre	La [34]	2009	OS DFS LRC	SUVMax MTV 50%	MTV50% (OS and DFS)	N/A	N/A	Increase of 17.4ml of MTV50% = HR 1.9 (first event) and 2.1 (death)
Pre	Machtay [17]	2009	OS DFS	SUVMax	SUVMax	9	2 year DFS 76 vs 37% (p=0.007) 2 year OS 82% vs 46 % (p=0.016)	DFS : 2.41 (p=0.03) OS : 2.47 (p=0.06)
Pre	Suzuki [62]	2009	OS DFS	SUVMax	None	5.5	N/A	N/A
Pre and Per	Farrag [15]	2010	OS DFS LRRFS	SUVMax	Pre : SUVMax (OS) Per SUVmax (OS)	Pre trt : 8.11 Per trt : 4.03	2-year OS SUVmax pre trt 81% vs 50% (p=0.027) SUVMax Per trt 82 vs 47% (p=0.026)	N/A
Pre and post	Moeller [44]	2010	DFS	SUVMax T Change in SUVMax T	SUVMax Post Change in SUVMax	6	N/A	N/A
Pre	Seol [63]	2010	DFS OS	SUVMax SUVMean MTV 2.5	MTV	9.3 cm^3	N/A	DFS : 2.19 (p=0.006) OS : 1.62 (p=0.051)

Pre	Deron [64]	2011	DFS OS	SUVMax MTV 50%	MTV50(DFS OS)	31 cm^3	N/A	N/A
Pre and per	Hentschel [39]	2011	DFS OS LRC	SUVmax SUVMean MTV 50%	ΔSUVmax10/20 (OS) MTV50 TEP 0 (OS)	ΔSUVmax10/20 50% MTV50% 10.2	2 year OS ΔSUVmax10/20 88% vs 38% (p =0.02) MTV 50% 83% vs 34% (p=0.02)	N/A
Pre and post	Hoshikawa [45]	2011	Recurrence	SUVMax	SUVMax Post % change in SUV	60%	N/A	Odds ratio local control 61.5 (p < 0.001)
Post	Murphy [49]	2011	DFS OS	SUVMax MTV 2, 2.5, 3, 3.5, 4 TLG	Post trt : MTV2.0 (DFS OS)	15 cm^3	N/A	Increase of 21cm^3 : 2.5 (DFS) and 2 (OS)
Pre	Schinagl [28]	2011	DFS OS	SUVMax SUVMean MTV 40% 50% MTV 2.5 GTV -PET	GTV PET (LC DFS OS in oral cavity and oropharyngeal cancer) MTV40% (DMFS DFS OS)	N/A	N/A	N/A
Pre per and post	Castaldi [21]	2012	RFS DSS	SUVMax Change in SUVMax (EORTC criteria)	EORTC criteria post	N/A	2 year DSS (late TEP) CR 100% PR 74% PD 33% p=.009	N/A
Pre	Chang [25]	2012	OS DFS LRFS	SUVMax MTV 2.5 TLG	TLG T (OS DFS)	65 g	3-year DFS 79.9% vs 37.4% (p<0.001)	DFS : 3.54 OS : 4.9
Pre	Chu [65]	2012	OS DFS	SUVMax MTV50% MTV Velocity (Difference between the 2 pre treatment TEP)	MTV T	N/A	N/A	Increase of 1cc/week = 85% increase of the risk of death
Pre	Higgins [22]	2012	DFS LRC OS	SUVMax SUVMean TLG (manuelly delineated)	SUVMean (DFS)	7 (median)	2 year DFS 82% vs 58% p=0.03	DFS : 1.14 (p=0.014)
Pre	Kao [27]	2012	DFS PRFS	MTV 2.5 3.0 40% 50%	MTV 2.5 (DFS PRFS)	13.6 ml	2-year PRFS 72% vs 39% (p=0.001) 2-year DFS 68% vs 41% (p=0.008)	DFS HR 2.69 p=0.011 PRFS HR 3.76 p=0.003
Pre	Romesser [66]	2012	OS DFS LRFS	SUVMax MTV (Gradient based method)	MTV	7.2	2-year LC 100 vs 54.2% (p=<0.001) DFS 94.7 vs 39.4% (p=0.001) OS 94.7 vs 64.2 (p=0.04)	N/A
Pre	Tang [32]	2012	OS DFS	SUVMax MTV50%	MTV50% T (OS and DFS)	Increase of 17cm^3 (difference between first and third quartiles)	N/A	DFS : 2.07 (p=0.00017) OS : 1.99 (p=0.0048)

Pre	Cheng [38]	2013	PFS DSS OS	MTV 2.5 TLG Normalized gray-level co-occurence matrix Neighborhood gray-tone difference matrix	TLG (PFS DSS OS) Uniformity (PFS DSS OS)	TLG 121.9 g Uniformity 4 bins 0.138	N/A	PFS TLG : 7.15(p=0.02) Uniformity : 0.32 (p=0.001) OS TLG : 5.85 (p=0.011) Uniformity : 0.46 (p=0.017)
Pre and post	Ashamalla [37]	2014	OS	SUVMax SUVMean Anatomical biological value = SUVMax X greatest tumor diameter	None	N/A	N/A	N/A
Pre and per	Chen [23]	2014	OS DFS	SUVMax pre et per (T et N) SUV reduction ratio	SUV reduction ratio tumor	3.9	2-year DFS 64% vs 41% (p=0.045) 2 year OS 66% vs 47% (p=0.035)	DFS : 2.33 OS : 2.64
Pre	Hanamoto [67]	2014	LR	SUVMax SUVMean MTV 2.5 TLG	For laryngeal and hypopharyngeal cancer, High MTV (> 25 ml) or high TLG (>144.8g) = high risk of partial response	N/A	N/A	13.4 (p=0.003)
Post	Ito [46]	2014	OS LC	SUVMax	SUVMax (OS)	6.1	OS 12.1 vs 44.6 months (p<0.001)	N/A
Pre	Romesser [26]	2014	LRC PFS OS	SUVMax MTV 42%	MTV (Distant metastasis, Disease progression or death)	9.7	5-year PFS 80.3% vs 56.7% (p=0.015) 5-year OS 84.1% vs 57.9% (p=0.008)	PFS : HR2.17 OS : HR 2.37
Pre	Sager [68]	2014	DFS OS	SUVMax MTV 50%	MTV50%	N/A	N/A	DFS : 2.5 OS : 2
Pre	Akagunduz [69]	2015	LRFS DFS OS	SUVMax SULMax MTV (adaptive threshold based)	MTV (treatment response, LR, Disease related death) SULMax (LR)	N/A	3-year (MTV) DFS 75.5% vs 25.3% OS 82.9% vs 55.9%	N/A
Pre	Cheng [29]	2015	PFS DSS	MTV 50 % 42% 2.5 and adaptive threshold TLG Grey level run lenght encoding matrix Grey level size zone matrix	Zone size nonuniformity, Uniformity, TLG adaptive threshold (PFS)	N/A	N/A	N/A
Pre	Hofheinz [33]	2015	PFS OS	SUVMax SUVMean MTV (adaptive threshold) TLG Aphericity	TLG MTV ASP	MTV 12.6 TLG 82.6 ASP 22%	N/A	PFS MTV : 2.89 (p=0.017) TLG : 3.11 (p=0.02) ASP : 3.09 (p=0.015) OS MTV : 3.3 (p=0.018) TLG : 3.32 (p=0.016) ASP : 5.9 (p=0.001)

Pre	Lin [30]	2015	Nodal relapse free survival OS DFS	SUVMax Nodal MTV2.5 N MTV40% N MTV50% N TLG40% N TLG50% N	TLG 40% (DFS NRFS)	38g	N/A	DFS : 2.12 (p=0.02)
Pre and post	Matoba [24]	2015	LRC PFS OS	SUVMax EORTC Criteria	EORTC criteria (OS and PFS)	N/A	N/A	N/A
Pre and per	Min [41]	2015	LRFS DFS MFFS OS	SUVMax MTV 2.5 TLG Percentage reduction between per and pre treatment PET	SUVMax per trt (DFS) MTV Pertrt (DFS) TLG pertrt (LRFS DFS)	SUVMax 4.25 MTV 3.3 TLG 9.4	N/A	N/A
Pre	Moon [70]	2015	DFS	SUVMax SUVMean MTV (adaptive threshold) TLG	TLG	7.6	N/A	DFS : 7.62 (p<0.001)
Pre	Rasmussen [18]	2015	Time to failure	SUVMax SUVMean SUVPeak	SUVmax	N/A	N/A	Time to failure (for SUVMax increase from 25th to 75th percentile) : 1.34 (p=0.039)
Pre and post	Schwartz [50]	2015	LR PFS OS	SUVMax SUVPeak MTV 40 %	Primary MTV 40% (LRR, DM, DFS)	8.76 cm^3	N/A	LRR : 4.01 (p=0.02) PFS : 2.34 (p=0.05)
Pre	Yabuki	2015	DFS OS	MTV 2 2.5 3 (for t and n)	MTV T 2.5 (OS DFS)	4.9 ml	3-year DFS 92.9% vs 38.6% (p<0.001) 3-year OS 95.35% vs 59.27% (p<0.001)	DFS : 6.97 (p=0.001) OS : 1.96 (p=0.002)
Post	Kim [47]	2016	PFS OS	SUVMax	SUVMax (DFS and OS)	4.4	3-year PFS 81.1% vs 42.9% 3-year OS 87.7% vs 56.9%	PFS : 4.79 (p<0.001) OS : 4.25 (p = 0.005)
Pre and per	Min [40]	2016	LRFS DFS MFFS OS	SUVMax MTV 2.5 TLG Percentage reduction between per and pre treatment PET	TLG pertrt (DFS) SUVMax per (DFS MFFS) MTV per (DFS OS) TLG per (DFS)	TLG per trt 9.4 SUVMax pre 11.45 and per 4.25 MTV pre 21.95 and per 3.3 TLG pre 91.75 and per 9.4	TLG 2-year DFS 85.9% vs 60.8% (p=0.005) MTV 2-year DFS 83.2% vs 62.3% (p=0.018) SUVMax 2-year DFS 82% vs 64.5% (p=0.025)	TLG DFS : 7.7 MTV DFS : 4.29 SUVMax : 4.18

Table 4: summary of the results of the 42 studies which analyzed the predictive value of PET before treatment. * : if considering only studies without volumetric PET parameters. DFS : Disease Free Survival, OS : Overall Survival, MTV : Metabolic Tumor Volume, TLG : Total Lesion Glycolysis

Quantitative parameters	Correlation with DFS/OS	Number of positive studies/total studies	Strength	Weakness
SUV_{Max}	Poorly	14/38 (11/14*)	Ease of use	Poorly reproducible No data concerning heterogeneity
SUV_{Peak}	?	0/2	More robust than SUV_{Max}	May not be representative of nonhomogeneous overall tumor uptake Ideal size of the ROI is still unclear
SUV_{Mean}	No	1/9	-	-
MTV/TLG	Yes	26/26	Represent the heterogeneity of the tumor uptake Ease of use	No clearly segmentation method No data concerning spatial relationships
Shape/ Texture analysis	?	3/3	Represent the heterogeneity of the tumor	No standardized method/Experimental Which correlation with histology ?