Positive and Negative Predictive Value of PET-CT in Skull Base Lesions: Case Series and Systematic Literature Review

John Peyton Hines1 Brittany E. Howard1 Joseph M. Hoxworth2 Devyani Lal1

1 Department of Otolaryngology–Head and Neck Surgery, Mayo Clinic, Phoenix, Arizona, United States 2 Department of Radiology, Mayo Clinic, Phoenix, Arizona, United States

Address for correspondence Devyani Lal, MD, Consultant, Department of Otolaryngology-Head and Neck Surgery, Mayo Clinic, 5777 E. Mayo Blvd., Phoenix, AZ 85054, United States (e-mail: lal.devyani@mayo.edu).


Abstract

Objectives To study positive (PPV) and negative predictive value (NPV) of positron emission tomography with computed tomography (PET-CT) scans in determining malignancy in skull base lesions and perform a systematic literature review for optimal PET-CT interpretation.

Design Retrospective case series and systematic literature review of the current English literature.

Setting Tertiary referral academic medical center.

Participants All patients with skull base lesions that underwent PET-CT and tissue biopsy from 2010 to 2013.

Main Outcome Measures PPV and NPV of radiologist's report and standardized uptake value (SUV) cutoff of 2.5 and 3, biopsy with pathologic interpretation, clinical follow-up.

Results A total of 31 PET-CT scans of 16 patients were studied; 10 PET-CT were performed upfront for diagnostic purposes and 21 were post-treatment surveillance scans. The PPV of radiologist's interpretation, SUV cutoff of 2.5, and SUV cutoff of 3.0 was 80%, 60%, and 68.4%, with a NPV of 100%, 83.3%, and 75%, respectively. Literature search yielded 500 abstracts; 7 studies met inclusion criteria for detailed review. No consensus or guidelines for optimal SUV cutoff value was found.

Conclusions PET-CT based on SUV cutoff criteria alone has high NPV but low PPV in determining malignancy in skull base lesions. Interpretation by a radiologist experienced in nuclear medicine and neuroradiology, synthesizing clinical, SUV, and radiologic data are of superior value.

Introduction

The skull base can be afflicted by a variety of inflammatory and neoplastic processes, and establishing accurate diagnosis is, therefore, critical to management. Accurate diagnosis of skull base lesions is challenging. This region is difficult to access surgically and is extensively traversed by critical neurovascular structures. Advances in diagnostic imaging have improved our ability to narrow the differential diagnosis and characterize lesions that may be malignant. No single imaging modality appears to be ideally suited to imaging the skull base. A complementary slew of tests such as magnetic resonance imaging (MRI), computed tomography (CT), radionuclide bone scan (Gallium or Technetium labeled), and Positron Emission Tomography scan with Computed Tomography (PET-CT) are therefore frequently utilized.1 PET-CT has been widely used in determining whether a lesion in the cranial base is malignant, both for

Keywords

► PET-CT ► skull base ► SUV


License terms
initial diagnostic workup and subsequent surveillance.\(^2\) In the absence of tissue-based diagnosis, the positive predictive value (PPV) and negative predictive value (NPV) of these radiologic tests become critical to understand. Currently, the criteria for interpreting a “positive” PET-CT at the skull base is not well defined.\(^3\) Several unique factors can affect the PPV of the PET-CT in skull base lesions. False positive results can arise from physiologic uptake of the brain, lymphoid tissue, salivary glands, and muscle. Additional non-malignant causes of positive results include infection, skull base osteomyelitis, inflammatory non-neoplastic disease, osteoradionecrosis, and post-radiation changes.\(^4,5\) Although PET-CT findings can be corroborated utilizing clinical observation, and other radiologic tests (such as MRI, SPECT/CT, or other forms of PET), the gold standard for confirming diagnosis are results from the tissue biopsy.

In recent years, significant advancements in minimal access surgical techniques have made many previously inaccessible skull base regions amenable to transnasal endoscopic or percutaneous imaging-guided biopsies. By comparing results from tissue biopsies and various imaging modalities, one can also better estimate their PPV and NPV in determining the presence of malignancy.

Often, a “positive” PET-CT cannot definitively establish whether a lesion is inflammatory or neoplastic, particularly in the setting of prior radiation therapy or surgery.\(^1\) The standardized uptake value (SUV), is a common benchmark to compare hypermetabolic tissues by measuring uptake of radiolabeled glucose in PET-CT.\(^2,5\) Although SUVs of greater than 2.5 have been used to indicate malignancy at other sites, an SUV cutoff value that optimizes PPV and NPV for distinction of skull base lesions has not been established.\(^6\) Moreover, a “normal” SUV value for the skull base has not been specifically determined, and few guidelines have been proposed to interpret SUVs.\(^7\) In part, this ambiguity stems from the fact that SUV is not an absolute value. Instead, it varies as a function of scanner, administered radiotracer dose, and patient body habitus / physiology. Improved precision in interpretation of PET-CT results is imperative, as significant consequences to clinical care may result from a false positive or false negative PET-CT.

The purpose of our study was to determine the PPV and NPV of the PET-CT in skull base lesions through a retrospective case series and a systematic review of the English literature.

### Methods

This study was approved by the institutional review board. A retrospective chart review of all patients with skull base lesions that underwent 18-fluoro-2-deoxyglucose (FDG) PET-CT as well as a tissue biopsy at our institution between 2010 and 2013 was conducted. At our institution, PET-CT is routinely employed in primary diagnostic work-up and staging of skull base lesions concerning malignancy as well as in post-treatment oncologic surveillance. Patients were included if they had at least one biopsy on initial evaluation, and also had subsequent clinical follow-up. Patients were excluded if they were not biopsied on initial evaluation or if they were lost to follow up after PET-CT interpretation. Data were collected from the electronic medical record including patient demographics, pathology, PET-CT findings, endoscopic exam, oncologic treatment, clinical follow-up, and oncologic outcomes. Using the tissue biopsy result as gold-standard, we compared the sensitivity, specificity, PPV, and NPV of PET-CT based on the following three criteria: the report of the reading radiologist (all of which were board-certified in both diagnostic radiology and nuclear medicine) not blinded to the patient’s history, SUV cutoff of 2.5, and SUV cutoff of 3.

A systematic review of the English literature from 1999 through August 2015 was conducted using PubMed and Embase databases. The search terms included: “PET, PET-CT, skull base, cranial base, nasopharynx, sinonasal, temporal, SUV, FDG, malignancy, neoplasm, osteoradionecrosis, and osteomyelitis.” The resultant abstracts were then screened by two authors (B. E. H. and J. P. H.). Next, all manuscripts from the selected abstracts were reviewed by three authors (J. P. H., D. L., and B. E. H.). Case reports, feasibility studies, and nonhuman studies were excluded. All selected manuscripts were reviewed and graded on level of evidence ranging from levels 1 to 5 (Oxford Center for Evidence-Based Medicine). Only studies with clearly established criteria for “positive” PET result, correlation with tissue biopsy or clinical follow-up, and discussion of PET result accuracy were considered.

### Results

#### Case Series

Thirty-one FDG PET-CT scans from 16 patients (11 males, 5 females) with lesions primarily of the skull base were identified (Table 1). Of these, 10 scans were performed for initial staging.

<table>
<thead>
<tr>
<th>PET-CT Characteristics</th>
<th>PET-CT (n = 31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease status</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>15</td>
</tr>
<tr>
<td>Negative</td>
<td>16</td>
</tr>
<tr>
<td>Indication</td>
<td></td>
</tr>
<tr>
<td>Initial staging</td>
<td>10</td>
</tr>
<tr>
<td>Surveillance imaging</td>
<td>21</td>
</tr>
<tr>
<td>Location of lesion</td>
<td></td>
</tr>
<tr>
<td>Anterior SB</td>
<td>20</td>
</tr>
<tr>
<td>Nasopharynx</td>
<td>8</td>
</tr>
<tr>
<td>Orbit</td>
<td>2</td>
</tr>
<tr>
<td>Clivus</td>
<td>1</td>
</tr>
<tr>
<td>Initial diagnosis</td>
<td></td>
</tr>
<tr>
<td>SNUC</td>
<td>5</td>
</tr>
<tr>
<td>Melanoma</td>
<td>3</td>
</tr>
<tr>
<td>NPC</td>
<td>2</td>
</tr>
<tr>
<td>Metastatic renal cell carcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Lymphoma anterior SB</td>
<td>1</td>
</tr>
<tr>
<td>Adenocarcinoma anterior SB</td>
<td>1</td>
</tr>
<tr>
<td>IgG4 Disease clivus</td>
<td>1</td>
</tr>
<tr>
<td>SCC anterior SB</td>
<td>1</td>
</tr>
</tbody>
</table>

Abbreviations: NPC, nasopharyngeal carcinoma; PET-CT, Positron Emission Tomography scan with Computed Tomography; SB, skull base; SCC, squamous cell carcinoma; SNUC, sinonasal undifferentiated carcinoma.
and metastatic workup of biopsy-positive malignancy prior to treatment and 21 for posttreatment surveillance. All 16 patients had a tissue biopsy performed at some point in their care. The median age of patients was 68.5 (range 50–91 years). Indications for PET-CT imaging were sinonasal undifferentiated carcinoma (n = 5), cutaneous/sinonasal melanoma (n = 3), nasopharyngeal carcinoma (n = 2), metastatic renal cell carcinoma (n = 2), sinonasal lymphoma of the anterior skull base (n = 1), sinonasal adenocarcinoma of the anterior skull base (n = 1), IgG4 disease of the clivus (n = 1), and sinonasal squamous cell carcinoma of the anterior skull base (n = 1). After the 31 PET-CT scans, 11 “negative” results were followed clinically without a biopsy and 20 PET-CT results reported to be “positive” by the reading radiologist and were subsequently biopsied except in one patient who was undergoing radiation therapy and had a cerebrovascular accident. Median length of followup was 10 months (0.5–33) with a mean length of 11.9 months (1–33) for negative results and 9.28 months (0.5–22) for positive results.

Of the 31 PET-CT exams performed, the initial radiologist interpretation was “positive” in 20 cases and “negative” in 11 cases. All patients interpreted as “positive” results had an SUV of at least 2.5. Of the 20 PET-CTs interpreted as “negative,” 3 had an SUV greater than or equal to 2.5. However, they were interpreted “negative” as the SUV value had decreased relative to prior exams. Of the 20 “positive” PET-CT results, there were only 4 false positives based on correlation with tissue histopathology. The reading radiologist interpretation had PPV of 80% and NPV of 100%, with sensitivity and specificity of 100 and 73.3%, respectively (Table 2). Eleven “negative” surveillance PET-CT results were confirmed to be true negatives by biopsy or a favorable evolution of the lesion on clinical followup. The mean SUV of truly positive PET-CT lesions was 9.36 (1.01–27). The mean SUV of falsely positive PET-CT lesions was 5.83 (3.1–9.3). The mean SUV of truly negative PET-CT lesions was 2.66 (1–5). False positive results in the skull base were secondary to posttreatment inflammation (n = 7), osteoradionecrosis (n = 2), and IgG4 disease (n = 1), corresponding to mean SUV max of 3.24, 5.8, and 9.3, respectively.

Next, SUV cutoff values of 2.5 and 3 were compared with benign from malignant lesions in the midline roof of the nasopharynx (MRN) by examining data from 4,846 Systematic Review
The literature search yielded 500 abstracts for review. A total of 101 manuscripts were selected for full manuscript review, but 94 of these had to be excluded due to lack of satisfactory description of qualitative or quantitative criteria used for reporting, lack of tissue biopsy or clinical follow-up, or lack of adequate statistical details or raw data. All seven studies included for further study were EBM Level 4. Authors differed widely in their criteria for “positive” PET-CT scan. "Positive" PET-CT was defined using SUV cutoff values in two studies, visual analog scale grading in one study, and intensity greater than surrounding tissue/physiologic independent of SUV in four studies (Table 3).

Chen et al sought to characterize PET uptake patterns in Waldeyer’s ring and refine criteria for defining “positive” lesions by reporting on a case series with both, prospective and retrospective arms. Of the 131 patients included, 110 patients were normal and 21 had newly-diagnosed nasopharyngeal carcinoma (NPC). Eighty of the normal patients had increased FDG uptake (defined as uptake intensity greater than physiologic liver uptake) in the lateral pharyngeal recess (LPR) of the nasopharynx. Using an SUV of 3.9 for cutoff of likely malignancy, sensitivity, and specificity were 72 and 80%, respectively. Applying N/P ratio (ratio of LPR uptake to palatine tonsil uptake) > 1.5 along with presence of symmetric uptake of LPR, normal or symmetric wall thickening, and detectable neck lymph node uptake were added to the SUV > 3.9 criteria for discriminating benign from malignant lesions, the sensitivity and specificity rose to 90.4 and 93.8%, respectively. Although sensitivity and specificity were higher when using both N/P ratio and SUV, there was no statistically significant difference apparent when each was used alone.

A second study by Chen et al focused on differentiating benign from malignant lesions in the midline roof of the nasopharynx (MRN) by examining data from 4,846

Table 2 Comparative PET-CT results

<table>
<thead>
<tr>
<th>Results</th>
<th>Radiologist Interpretation</th>
<th>Positive @ 2.5 SUV (n = 31)</th>
<th>Positive @ 3.0 SUV (n = 31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>True positives</td>
<td>16</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>False positive</td>
<td>4</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>True negatives</td>
<td>11</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>False negatives</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100%</td>
<td>93.8%</td>
<td>81.3%</td>
</tr>
<tr>
<td>Specificity</td>
<td>73.3%</td>
<td>33.3%</td>
<td>60%</td>
</tr>
<tr>
<td>NPV</td>
<td>100%</td>
<td>83.3%</td>
<td>75%</td>
</tr>
<tr>
<td>PPV</td>
<td>80%</td>
<td>60%</td>
<td>68.4%</td>
</tr>
</tbody>
</table>

Abbreviations: NPV, Negative Predictive Value; PPV, Positive Predictive Value; SUV, standardized uptake value.
participants enrolled in a FDG PET-CT cancer-screening program. A total of 121 participants were analyzed, including 25 patients with newly diagnosed NPC, 30 normal control participants, and 66 participants (39 symptomatic and 27 asymptomatic) with focally increased FDG uptake defined as uptake in the MRN greater than or equal to physiologic uptake in the liver. After calculating the ratio of uptake in the MRN/palatine tonsil (MRN/PT), they measured sensitivity and specificity employing different SUV cutoff criteria. Using an MRN SUV > 4.61 and MRN/PT > 1.14 along with presence of symmetric uptake of LPR, normal or symmetric wall thickening, detectable neck lymph node uptake, and characteristic axis of MRN measurements allowed for improved discrimination of benign from malignant disease with a sensitivity and specificity of 95 and 85.5%, respectively.

Yen et al8 prospectively examined 27 patients with both 201TI (thallium-201) single photon emission computed tomography (SPECT/CT) and FDG PET-CT to compare the accuracy of these imaging modalities for detection of recurrent nasopharyngeal malignancies. All patients had NPC treated with radiation therapy and chemotherapy, and had followup MRI suggestive of local recurrence or metastasis. PET-CT “positivity” was determined using a fourpoint visual analog scale (one representing no abnormal uptake and four representing intense uptake, greater than that of background tissue). Scans were interpreted as “negative” if graded one or two, and “positive” for recurrence if graded three or four. The PPV, NPV, sensitivity, and specificity of FDG PET-CT were 81.3, 81.8, 86.7, and 75%; in comparison to PPV, NPV, sensitivity, and specificity of 100, 70.6, 66.7, and 70.6% for SPECT/CT, respectively. This study suggested that SPECT/CT is a better imaging modality than PET-CT for intracranial recurrence because of the high intracranial lesion/background ratio of 201TI. Although SPECT/CT has less false positives than PET-CT, it may be more likely to miss the detection of small recurrent lesions (< 1.5 cm).

Cheon et al5 retrospectively reviewed the records of 45 patients with previously-treated head and neck cancer, 13 of which were NPC. They sought to compare whole-body FDG-PET with conventional imaging modalities such as CT and MRI in the detection of disease recurrence. Lesions were considered “positive” if FDG uptake was increased focally in the lesion and was greater than surrounding tissue or SUV > 3. PET was more accurate than conventional imaging modalities in detection of cancer recurrence, especially within the first 6 months after treatment. When considering nasopharyngeal location alone, PET had a PPV of 100% and NPV of 88.9% with a sensitivity and specificity of 80 and 100%, respectively. For all sites, PET scans were found to be more accurate in determining recurrence after surgery than after radiotherapy.

Kao et al10 prospectively followed 36 patients with NPC treated with primary radiotherapy and sought to evaluate
PET-CT Criteria in Skull Base Lesions  Hines et al.

Discussion

The skull base is challenging to evaluate with conventional imaging. Prior radiotherapy or surgical intervention can further confound evaluation of this problematic area. FDG PET is an imaging modality that allows characterization of lesions based on their glycolytic activity. The high metabolic uptake of this radiolabeled glucose analog in tumor cells allows differentiation of malignant from benign processes, although spatial resolution and anatomical landmarks are limited.

Studies have shown that by integrating FDG PET data with a concurrent CT, delineation of viable tumor is superior to PET, CT, or MRI alone. Due to its established high sensitivity and NPV, FDG PET-CT has a role in the diagnosis, staging, and followup of skull base malignancies and is frequently utilized in the workup and management of such lesions. However, FDG uptake is not specific to tumors as other types of high metabolic activity may show increased uptake, (i.e., inflammation, infection, and physiologic uptake of brain, salivary gland, muscle, and lymphoid tissue) resulting in a low specificity and PPV. Utilization of SUV alone as a marker of malignancy has been studied extensively in the evaluation of head and neck cancer with no definite consensus on optimal cutoff values to maximize the PPV and NPV of PET-CT. Some authors suggest that SUV levels are not a distinguishing factor for determining true positivity, and that visual pattern interpretation was more important. There have been limited viable options to increase PPV of PET-CT without significantly decreasing its NPV. Therefore, there is the need to define clinically and cost effective criteria for positive PET-CT, as well as to look at viable alternative options for imaging.

Alternative imaging options for the skull base include SPECT, other forms of PET (11C-choline, 2-18F-fluoro-L-tyrosine, fluoroethyltyrosin), MRI, and PET/MRI. Some authors propose that SPECT/CT is better than FDG PET-CT for intracranial recurrence and may also help clarify false positive FDG uptake caused by osteoradionecrosis because 201Tl has a high intracranial lesion/background ratio with less accumulation within inflammatory or necrotic tissues but it may be more apt to missing small recurrent lesions. There are studies suggesting that 11C-choline PET-CT could be a preferable modality for skull base imaging because of its low background radioactivity in normal brain tissue, which may allow improved detection of intracranial, skull base, and orbital invasion. It may be more suitable for patients with diabetes mellitus because hyperglycemia does not influence 11C-choline PET-CT results. Tomura et al compared 11C-choline PET-CT with FDG PET-CT and concluded that 11C-choline PET-CT has superior potential in skull base imaging due to superior inter-rater agreement and higher uptake for tumors at the skull base, but 11C-choline has higher uptake in mucosa and in regions of inflammation. Others have investigated fluoroethyltyrosin-PET (PET/FET) for its value in skull base imaging. In employing PET FET along with MRI after proton beam therapy for skull base tumors, Korchi et al described patterns indicative of radionecrosis. Using PET PET mean tumor-to-background ratio < 1.95 or a pattern of increasing uptake over time, they suggested that PET PET may be a useful adjunct for differentiation of radiation necrosis from tumor recurrence.

Others advocate that MRI continues to be the imaging modality of choice in assessing intracranial or dural involvement of malignancy and indicate that FDG PET-CT underestimates tumor invasion of the skull base in comparison to MRI. In comparison to MRI, Teo et al report that FDG PET is more effective in detecting recurrent/residual NPC and FDG PET-CT identifies viable tumor better in inflammatory...
areas after irradiation, but MRI does show a trend toward higher accuracy in disease detection at the primary site along with advantages in detecting intracranial lesions, adenopathy, and metastases. Integrated PET-MRI systems have become commercially available in recent years, though the role of this new technology in oncologic imaging is still being defined. Initial studies that evaluated PET-MRI in the context of head and neck cancer have confirmed feasibility and shown good diagnostic capability. The improved resolution and tissue contrast afforded by PET-MRI should be particularly useful in evaluating the skull base. However, none of the PET-MRI studies to date have specifically evaluated this area. Future studies will need to determine if PET-MRI can supplant PET-CT for the evaluation of skull base neoplasms and, if so, whether different SUV thresholds need to be employed.

Reviewing the results of this and previously published studies, it is evident that the optimal means of identifying a PET-CT positive result has not been clearly established.

In this study, the radiologist’s report had the most favorable PPV and NPV for viable malignancy at the skull base as opposed to standard SUV cutoff values. The radiologist’s approach to image interpretation utilized SUV, pattern recognition, and prior exam comparison, and resulted in the highest PPV and NPV (80% and 100%, respectively) in our series.

Our study had several limitations. There were small number of patients (n = 16) in the series due to our inclusion criteria and relative infrequency of skull base lesions. In regards to the type of disease and inclusion of both pre- and posttreatment PET-CT, there was significant heterogeneity in our study group as well as in the systematic review. Additionally, the dominance of NPC comparative to other skull base pathology in the current literature limits the conclusions that can be drawn regarding diverse skull base lesions. These factors led to difficulty in comparing and generalizing due to differences in subsites, pathology and biologic uptake rate of FDG, prior diagnostic and treatment modalities, and ultimately different management of patients. Variations in technique and timing of PET may also lead to differences in the SUV. Although we chose tissue biopsy as gold standard, sampling errors and errors in the pathologists’ interpretation may also result in false negative results. Ultimately, our goal was not to define the accuracy of PET-CT in a specific disease, but to document our experience and review the literature to highlight the challenges and limitations associated with interpreting PET-CT results in a diverse group of patients with skull base disease where there are no well-defined thresholds for distinguishing benign from malignant disease. Future studies, including a larger patient population and well-defined criteria for “positivity,” are needed to confirm our findings and elucidate what imaging modality and cutoff would be most efficacious in the skull base. The cost-effectiveness of FDG PET-CT in the skull base imaging has yet to be completely determined. Some believe that integrating PET and CT in one scan allows for improved evaluation, which may be more cost-effective than independent imaging tests. However, the costs of further workup and additional morbidity to patients with a “positive” PET-CT should also be considered.

Conclusions

No consensus exists with regard to ideal SUV cutoff values in interpreting PET-CT scans of the skull base. While the high NPV of a “negative” PET-CT report has a role in ruling out skull base malignancy, a “positive” result warrants cautious interpretation. Utilizing “standard” SUV cutoff values of 2.5 or 3 to define positivity can result in a high number of false positive results. This may lead to additional, potentially unnecessary interventions, and morbidity to patients. Our case series found that the report of a radiologist experienced in both, neuroradiology and nuclear medicine is superior to a SUV cutoff value for both PPV and NPV. This interpretation utilizes a nuanced study of the pattern and character of FDG uptake in conjunction with clinical data and previous imaging. Although tissue confirmation is ideal for lesions with a high FDG uptake, procedural morbidity should be weighed against vigilant clinical observation if conflicting clinical evidence is present.

Notes

EBM Level: 3
This article was presented at the 24th Annual North American Skull Base Society Meeting, February 2014, San Diego, California, United States.

Funding Disclosure
None.

Conflict of Interest
None.

References

5 Ramakrishnan VR, Lee JY, O’Malley BW Jr, Palmer JN, Chiu AG. 18F-FDG-PET in the initial staging of sinonasal malignancy. Laryngoscope 2013; 123(12):2962–2966


Tomura N, Mizuno Y, Saginoya T. PET/CT findings for tumors in the base of the skull: comparison of 18 F-FDG with 11 C-methionine. Acta Radiol 2015; [Epub ahead of print].