Diagnostic Accuracy of 18F-FDG-PET in Patients with Testicular Cancer: a Meta-analysis

Jing-Yi Zhao, Xue-Lei Ma, Yan-Yan Li, Bing-Lan Zhang, Min-Min Li, Xue-Lei Ma, Lei Liu

Abstract

Objective: Fluorine-18-fluorodeoxyglucose positron emission tomography (18F-FDG-PET) is a new technique for identifying different malignant tumors using different uptake values between tumor cells and normal tissues. Here we assessed the diagnostic accuracy of 18F-FDG-PET in patients with testicular cancer by pooling data of existing trials in a meta-analysis. Methods: PubMed/MEDLINE, Embase and Cochrane Central Trials databases were searched and studies published in English relating to the diagnostic value of FDG-PET for testicular cancer were collected. The summary receiver operating characteristic (SROC) curve was used to examine the FDG-PET accuracy. Results: A total of 16 studies which included 957 examinations in 807 patients (median age, 31.1 years) were analyzed. A meta-analysis was performed to combine the sensitivity and specificity and their 95% confidence intervals (CIs), from diagnostic odds ratio (DOR), positive likelihood ratios (PLR), negative likelihood ratio (NLR). SROC were derived to demonstrate the diagnostic accuracy of FDG-PET for testicular cancer. The pooled sensitivity and specificity were 0.75 (95% confidence interval (CI), 0.70-0.80) and 0.87 (95% CI, 0.84-0.89), respectively. The pooled DOR was 35.6 (95% CI, 12.9-98.3). The area under the curve (AUC) was 0.88. The pooled PLR and pooled NLR were 7.80 (95% CI, 3.73-16.3) and 0.31 (95% CI, 0.23-0.43), respectively. Conclusion: In patients with testicular cancer, 18F-FDG-PET demonstrated a high SROC area, and could be a potentially useful tool if combined with other imaging methods such as MRI and CT. Nevertheless, the literature focusing on the use of 18F-FDG-PET in this setting still remains limited.

Keywords: 18F-FDG-PET - testicular cancer – diagnosis - meta-analysis

Cancer Center and State Key Laboratory of Biotherapy, West China Hospital, West China Medical School, Sichuan University, Chengdu, China *Equal contributors *For correspondence: mxlmax@163.com, liuleihx@gmail.com

Introduction

Testicular cancer is the most frequently diagnosed malignancy in the age group of 20 to 40 years (Devesa et al., 1995), which is traditionally subdivided into seminomatous and nonseminomatous tumors. The worldwide incidence ranges from 0.5 to 9.9 cases per 100,000 men per year, with extensive regional differences. And the most important risk factor remains cryptorchism and there is a variable association with testicular microlithiasis (Becherer, 2011). Before the use of cisplatin, the cure rates were less than 10% (Ganjoo et al., 1999). Today, testicular cancer is highly curable, with 5-year survival rates over 95% (Lewis et al., 2006). More than 90% of patients are cured with surgery, radiotherapy, and chemotherapy alone or combination of them (Yetişyigit et al., 2014). This success depends on the accuracy of disease diagnosis and the application of optimum treatment. Because of the excellent cure rates in testicular cancer, new diagnostic methods are needed to strengthen the diagnostic value for early diagnosis and correct staging to ensure that the optimum treatment strategy is employed.

In order to correctly evaluate patients with testicular cancer, a variety of diagnostic imaging modalities have been used. Computer tomography (CT) and magnetic resonance imaging (MRI) are the common imaging modalities in testicular cancer with the role in initial staging of patients, in assessment of cancer response and detection of disease relapse. However, dependence on structural abnormal conditions to identify disease means CT and MRI have many limitations. It cannot detect disease in “normal-sized” lymph nodes or cannot detect whether there is active tumor in residual masses after chemotherapy (Nichols, 1998; Huddart, 2003). As well as chest X-ray and ultrasound of the abdomen, they also have the same limitation. Tumor maker is another diagnostic sign, however, it also has limitations, because there is no serum tumor marker in the seminomatous tumors, only nonseminomatous tumor has tumor marker (Becherer, 2011); furthermore, tumor makers always have unsatisfied sensitivity and specificity. So, it is being active to explore new diagnostic method, in particular functional imaging, such as [18F] -fluorodeoxyglucose positron emission tomography, has significant advantage over anatomically...
Fluorine-18-fluorodeoxyglucose positron emission tomography (18F-FDG-PET) has been successfully used to evaluate different types of malignant tumors (Saunders et al., 1999; Czernin, 2002; Halfpenny et al., 2002; Johns Putra et al., 2004; Rohren et al., 2004; Chen et al., 2011), like esophagus, thyroid, nasopharyngeal and lung carcinoma (Mutlu et al., 2013; Uzel et al., 2013). FDG-PET is a non-invasive functional imaging method in this setting.

The accuracy of FDG-PET for diagnosis of testicular cancer has been assessed in a series of studies. However, a meta-analysis of published data in this field is lacking. The purpose of this study is to meta-analyze published data on the diagnostic performance of FDG-PET in patients with testicular cancer and to assess the overall accuracy of the functional imaging method in this setting.

Materials and Methods

Search strategy and study selection

A comprehensive computer literature search of the PubMed/MEDLINE, Embase and Cochrane Central Trials databases was conducted to find relevant published articles on the diagnostic accuracy of FDG-PET in patients with testicular cancer. We used the following search terms: “testicular cancer” and “PET”. The search was performed in consultation with a third reviewer after face-to-face discussion.

Studies were excluded if they met the following criteria: (1) did not evaluate testicular cancer; (2) incomplete data available; (3) were duplicated or updated; (4) reviews, editorials, corresponding letters that did not report their own data; and (5) case reports.

Two independent reviewers (JY Zhao and XL Ma) reviewed the titles and abstracts of the retrieved articles, applying the selection criteria mentioned above. Articles were excluded if they were clearly ineligible. The same two researchers then independently reviewed the full-text version of the remaining articles to determine their eligibility for inclusion. Disagreements were resolved in consultation with a third reviewer after face-to-face discussion.

Statistical methods

Study-level analysis: Based on the value of true-positive (TP), true-negative (TN), false positive (FP), false negative (FN), we calculated the pooled sensitivity and specificity, positive and negative likelihood ratios (LR) and diagnostic odds ratio (DOR). The calculated statistics above were used to examine the FDG-PET accuracy for the diagnosis of testicular cancer. All statistics were reported as point values with 95% confidence intervals (CIs). Sensitivity was defined as the TP rates and calculated as TP/ (TP+FN). Specificity was defined and calculated as TN/ (FP+TN). LR indicates how much use of a given test would raise or decrease the probability of having disease. In this study, the PLR was the measure of the likelihood that a positive staging result of an index test would occur in a patient with testicular cancer, while the NLR was the measure of the likelihood that a negative staging result would occur in a patient without testicular cancer. The diagnostic odds ratio (DOR) is a single overall indicator of diagnostic performance and expresses how much greater the odds of having the disease for the people with a positive test result than for the people with a negative test result. The DOR was calculated as (TP x TN) / (FP x FN). In addition, summary receiver operator characteristics (SROC) curves were constructed to examine the interaction between sensitivity and specificity.

We use the area under the curve (AUC) to measure the overall performance of the diagnostic test (Moses et al., 1993). Statistical analyses were performed using Metadisc statistical software version 1.4 and Stata software version 11.1.

Meta-analysis model: The between-study heterogeneity was evaluated by computing Higgins’s I² and X² tests for heterogeneity using the generic inverse variance method of meta-analysis. A random effects model is used for statistical pooling of the data in the case of heterogeneity between the studies (P<0.1); a fixed effects model for statistical pooling of the data is used if there was no heterogeneity between the studies.

Quality of studies: We assessed the quality of the included studies in this meta-analysis using a checklist.
Table 1. Main Characteristics of All Studies Included in the Meta-analysis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Country</th>
<th>No.of patients</th>
<th>No.of examinations</th>
<th>Age (year)</th>
<th>PET technique</th>
<th>Reference</th>
<th>QUADAS score</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.de Wit, 2008</td>
<td>Germany</td>
<td>72</td>
<td>1</td>
<td>28</td>
<td>18F-FDG 388 MBq 60 min uptake histo</td>
<td>histopathologic</td>
<td>14</td>
<td>21</td>
<td>1</td>
<td>11</td>
<td>39</td>
</tr>
<tr>
<td>M.Bachner, 2012</td>
<td>Europe, Israel</td>
<td>125</td>
<td>3</td>
<td>NA</td>
<td>18F-FDG, histopathologic, follow-up data</td>
<td>histopathologic, follow-up data</td>
<td>10</td>
<td>14</td>
<td>8</td>
<td>7</td>
<td>41</td>
</tr>
<tr>
<td>M.de Santis, 2004</td>
<td>Austria, Germany</td>
<td>51</td>
<td>1</td>
<td>NA</td>
<td>18F-FDG 370 MBq 45 min uptake histo</td>
<td>histopathologic</td>
<td>11</td>
<td>22</td>
<td>3</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>K.Oechsle, 2008</td>
<td>Germany</td>
<td>121</td>
<td>3</td>
<td>30</td>
<td>18F-FDG 200-400 MBq 15 min uptake histo</td>
<td>histopathologic</td>
<td>10</td>
<td>16</td>
<td>2</td>
<td>2</td>
<td>41</td>
</tr>
<tr>
<td>P. Tsatalpas, 2002</td>
<td>Germany</td>
<td>23</td>
<td>2</td>
<td>NA</td>
<td>18F-FDG 100-390 MBq 50-90 min uptake histo</td>
<td>histopathologic, follow-up data</td>
<td>10</td>
<td>22</td>
<td>8</td>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>U. Cremerius, 1999</td>
<td>Germany</td>
<td>50</td>
<td>3</td>
<td>31</td>
<td>18F-FDG 159-283 MBq 30-60 min uptake histo</td>
<td>histopathologic, follow-up data</td>
<td>11</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>C.Bokemeyer, 2002</td>
<td>Germany</td>
<td>23</td>
<td>3</td>
<td>31</td>
<td>18F-FDG 250-500 MBq 45-60 min uptake histo</td>
<td>follow-up data</td>
<td>10</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>U. Cremerius, 1998</td>
<td>Germany</td>
<td>33</td>
<td>3</td>
<td>30</td>
<td>18F-FDG 120-360 MBq 40-80 min uptake histo</td>
<td>histopathologic, follow-up data</td>
<td>10</td>
<td>16</td>
<td>3</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>SF Hain, 2000</td>
<td>England</td>
<td>55</td>
<td>3</td>
<td>30</td>
<td>18F-FDG 185-370 MBq 45 min uptake histo</td>
<td>histopathologic</td>
<td>13</td>
<td>12</td>
<td>0</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>P. Albers, 1999</td>
<td>Germany</td>
<td>37</td>
<td>3</td>
<td>NA</td>
<td>18F-FDG 185-370 MBq 45 min uptake histo</td>
<td>histopathologic</td>
<td>13</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>A. Becherer, 2004</td>
<td>Austria, Germany</td>
<td>55</td>
<td>3</td>
<td>NA</td>
<td>18F-FDG 185-370 MBq 45 min uptake histo</td>
<td>histopathologic</td>
<td>13</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>U. Lassen, 2003</td>
<td>Denmark</td>
<td>46</td>
<td>3</td>
<td>NA</td>
<td>18F-FDG 200-220 MBq 60 min uptake histo</td>
<td>follow-up data</td>
<td>10</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>S. Hinz, 2008</td>
<td>Germany</td>
<td>20</td>
<td>4</td>
<td>42.5</td>
<td>18F-FDG 185-370 MBq 45 min uptake histo</td>
<td>histopathologic</td>
<td>13</td>
<td>14</td>
<td>2</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>J.R.Spermon, 2002</td>
<td>Netherland</td>
<td>50</td>
<td>2</td>
<td>30</td>
<td>18F-FDG 200-220 MBq 60 min uptake histo</td>
<td>histopathologic</td>
<td>10</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>J. Siekiera, 2012</td>
<td>Poland</td>
<td>37</td>
<td>2</td>
<td>NA</td>
<td>18F-FDG 185-370 MBq 45 min uptake histo</td>
<td>histopathologic</td>
<td>13</td>
<td>14</td>
<td>2</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>Z. Akbulut, 2011</td>
<td>Turkey</td>
<td>16</td>
<td>2</td>
<td>NA</td>
<td>18F-FDG 185-370 MBq 45 min uptake histo</td>
<td>histopathologic</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

NA, indicates not available; TP, true-positive; FP, false-positive; FN, false-negative; TN, true-negative

Based on the quality assessment for the studies of diagnostic accuracy (QUADAS) tool (Whiting et al., 2003). It has 14 items to evaluate the quality of studies. Each item may be scored “yes” if reported; “no” if not reported; or “unclear” if there is no adequate information in the article. The quality assessment was done by JY Zhao and XL Ma independently. Disagreements were resolved in consultation with a third reviewer after face-to-face discussion.

Publication bias: Publication bias was assessed using Begg’s funnel plots and test (p > 0.05 was considered that there was no potential publication bias) (Deeks et al., 2005). The Begg’s funnel plots and test were performed by Stata 11.0.

Results

Study Selection and Characteristics of studies in our analysis

The comprehensive computer literature search from the PubMed/MEDLINE, Embase, and Cochrane Central Trials databases revealed 179 articles, 37 potentially relevant studies were restricted for evaluation. Ultimately, 21 studies were excluded for undesirable article types (n=21), not written in English (n=6), and insufficient data (n=4), letter (n=2), case report (n=1), and review (n=8). Thus, 16 studies were included in our final dataset for the meta-analysis (Cremerius et al., 1998; Albers et al., 1999; Cremerius et al., 1999; Hain et al., 2000; Bokemeyer et al., 2002; Spermon et al., 2002; Tsatalpas et al., 2002; Lassen et al., 2003; De Santis et al., 2004; Becherer et al., 2005; de Wit et al., 2008; Hinz et al., 2008; Oechsle et al., 2008; Akbulut et al., 2011; Bachner et al., 2012; Siekiera et al., 2012). The flowchart of study selection was shown in Figure 1.

Table 1 summarized the main characteristics of the included studies. A total of 893 examinations in 807 patients (median age, 31.1 years) were evaluated. Histopathologic findings or follow-up data were used as the reference standard for the final result of PET for testicular tumor in nine studies. Five studies included patients who were only referred to histopathologic findings, and two studies referred to follow-up data (Table 1). Detailed information about the QUADAS

![Figure 1. Flow Chart of Selection Process for Eligible Studies](image-url)
The questionnaire of the selected studies is shown in Table 2. The overall quality of the studies was good, with most items rating as “yes” rating for nearly all items.

### Accuracy of PET for the diagnosis of testicular tumor

The meta-analysis was based on the extracted number of TP, TN, FP, and FN for the diagnosis of testicular cancer to assess the accuracy of FDG-PET in the included studies. The pooled sensitivity and specificity of FDG-PET was calculated using a random effects model. The pooled sensitivity and specificity for the testicular tumor were 0.75 (95% CI, 0.70-0.80) and 0.87 (95% CI, 0.84-0.89), respectively (Figure 2A).

The pooled PLR and pooled NLR were 7.80 (95% CI, 3.73-16.32) and 0.31 (95% CI, 0.23-0.43), respectively (Figure 2B). The pooled DOR was 35.57 (95% CI, 12.87-98.29, Figure 3), and the SROC was 0.88 (Figure 4). There were statistically significant heterogeneity in SPE ($X^2 \; P < 0.001, I^2 = 89.8\%$), PLR ($X^2 \; P < 0.001, I^2 = 89.7\%$) and DOR ($X^2 \; P = 0.001, I^2 = 78.5\%$), respectively. However, according to the meta-regression analysis, the accuracy of PET measurement was not affected by the covariates.

### Assessment of publication bias

We used Begg’s test and funnel plot to examine publication bias. Fortunately, there is no significant
Discussion

To our knowledge, this meta-analysis is the first to evaluate the diagnostic performance of FDG-PET in patients with testicular cancer. We know now testicular cancer most frequently diagnosis in young men. Due to excellent cure rates in testicular cancer, new diagnostic methods are needed to correct diagnosis which is particularly important to choose a most appropriate therapeutic schedule. Several studies have reported the usefulness of FDG-PET in the diagnosis of testicular cancer (Czernin, 2002; Huddart, 2003). However, one of the major problems with these studies is that many have limitations, for their analyzing only relatively small numbers of patient. To derive more strong evaluates of diagnostic performance of FDG-PET we pooled published studies using meta-analysis.

In this meta-analysis, we evaluated the accuracy of FDG-PET, a non-invasive technique, using functional imaging rather than structure imaging for diagnosis testicular tumor. The results of our meta-analysis demonstrate that FDG-PET had a high diagnostic accuracy for accessing testicular cancer with a summary AUROC of 0.88. According to this result, FDG-PET can be used in clinical practice as a good tool for the diagnosis of testicular cancer (Hanley and McNeil, 1982). The selection of a most appropriate treatment for testicular cancer is on the basis of early diagnosis and correct staging of the disease. Although biopsy is the gold standard for detection and characterization of testicular cancer, it is an invasive method and not fit for monitoring patients in clinical practice. In this respect, FDG-PET is promising and worthy to translate into clinical practice because it is a noninvasive and reliable method for the detection of testicular tumor. In addition, FDG-PET could be contributed to cancer staging by identifying lymphonodus that most likely to be malignant. The ease of practical and application, non-invasive character of FDG-PET make it increasingly popular among radiologist. In addition, FDG-PET demonstrated a good specificity, being a potentially useful tool if combined with other imaging methods.

However, this approach still has several limitations. Residual masses in the pelvis cause certain diagnostic problems due to adjacent urine activity. FDG-PET imaging is based on FDG uptake values; the uptake value in active tumor is higher than that in normal tissues. Excretion though urinary system and gather in bladder of FDG result in primary radiologist unable to distinguish the vicinal tumor with high uptake values. Sufficient hydration before FDG injection and emptying the bladder before the examination could overcome the problem. A second scan of the suspicious region after showing a wash-out bladder can also differentiate the tumor and urinary activity. Moreover, attenuation correction combined with iterative reconstruction algorithms help to differentiate small structures from bladder activity (Becherer et al., 2005).

Residual masses are larger than 3 cm in patients after chemotherapy is considered to be a problem for false-positive results (Lewis et al., 2006). And another possible reason for false-positive is nonspecific inflammatory processes after chemotherapy (Kollmannsberger et al., 2002; Tsatalpas et al., 2002; Albers et al., 2004; Johns Putra et al., 2004). In this regard, it might be of principal importance that we should do the examination before chemotherapy or have an interval of at least 4 weeks after chemotherapy (De Santis et al., 2004; Becherer et al., 2005). Moreover, other imaging methods, such as CT, MRI, combined with FDG-PET can help to avoid the false-positive findings. And on the other hand, close cooperation between the physicians and those who explain the false-positive findings. And on the other hand, close cooperation between the physicians and those who explain the false-positive findings.
of chemotherapy may result in false-negative (Cremerius et al., 1998, Hain, 2005). In order to avoid these false-negative results, a close cooperation between physicians and examiners are important. FDG-PET combined with other imaging methods can also help to overcome the problem. Moreover, another possible reason for false-negative would be considered. Testicular cancer is divided into two major groups: seminoma and nonseminomatous germ cell cancer. The latter group consists of not only embryonal carcinoma, teratoma, yolk sac tumors, but other rare tumor types. FDG-PET scans cannot distinguish fibrosis from differentiated teratoma as both are low uptake values (Stephens et al., 1996; Sanchez et al., 2002; Dalal et al., 2006; Lewis et al., 2006; Becherer, 2011). In this connection, an enormous number of researches are required to make a new better criterion to identify vital teratoma and nonvital tissues.

Publication bias is a major concern in all forms of pooled analyses, as studies reporting significant results are more likely to be published than those reporting non-significant results. Indeed, it is not unusual for early small-sized studies to report a positive relationship that subsequent larger studies fail to repeat. Fortunately, in this meta-analysis, no significant publication bias was found by the Begg’s test and funnel plot. The meta-regression did not show any relationship between the characteristics of studies and the diagnostic odds ratio.

Some limitations of this study should also be taken into consideration. First, many studies were excluded from the analysis, resulting in a small number of studies in the final meta-analysis. Second, the pooled results of the meta-analysis had high statistical heterogeneity. The heterogeneity has multiple sources, including differences in the reference test, difference in the study designs, differences in the subject populations, differences in cut-off values and differences in diagnostic modes (initial diagnosis or evaluation of residual masses after therapy). These could influence the stiffness value and lead to an overestimation of the true diagnostic performance.

In conclusion, 18F-FDG-PET is an accurate noninvasive and useful diagnostic tool for the patients with testicular cancer. FDG-PET is able to differentiate between nonvital and vital lesions in patients with testicular cancer. A negative PET eliminates viability in large lesions and contributes to avoid unnecessary surgery. On the other hand, a positive PET is a predictor of a viable tumor with the relapse risk. FDG-PET demonstrated a good specificity, being potentially useful tools if combined with other imaging methods such as MRI, CT. In addition, FDG-PET can provide uptake values, which can used as a prognostic factor in many tumors, and this is a new research hotspot. So, we also consider using it for predicting the prognosis depended on the difference of uptake values in tumor. In addition, planning of radiation therapy might become another application and research field for FDG-PET in the future.

Acknowledgements

The authors declare that they have no conflict of interest.

References


Spermon JR, De Geus-Oei LF, Kiemeney LA, Wijten JA & Oyen WJ (2002). The role of (18)fluoro-2-deoxyglucose positron emission tomography in initial staging and re-staging after chemotherapy for testicular germ cell tumours. BJU Int, 89, 549-56.


