Diagnostic value of PET with different radiotracers and MRI for recurrent glioma: a Bayesian network meta-analysis

Tian Xiaoxue,1 Wang Yinzhong,2 Qi Meng,3 Xingru Lu,2 Junqiang Lei2

ABSTRACT

Objective The purpose of this study was to evaluate the diagnostic accuracy of 6 different imaging modalities for differentiating glioma recurrence from postradiotherapy changes by performing a network meta-analysis (NMA) using direct comparison studies with 2 or more imaging techniques.

Data sources PubMed, Scopus, EMBASE, the Web of Science and the Cochrane Library were searched from inception to August 2021. The Confidence In Network Meta-Analysis (CINeMA) tool was used to evaluate the quality of the included studies with the criterion for study inclusion being direct comparison using 2 or more imaging modalities.

Main outcomes and measures The consistency was evaluated by examining the agreement between direct and indirect effects. NMA was performed and the surface under the cumulative ranking curve (SUCRA) values were obtained to calculate the probability of each imaging modality being the most effective diagnostic method. The CINeMA tool was used to evaluate the quality of the included studies.

INTRODUCTION

Glioma is one of the most common primary brain tumours and accounts for about 25% of primary tumours in the central nervous system.1 The treatment of glioma is very difficult, especially for the high-grade glioma such as glioblastoma multiforme (GBM). Even with the standardised treatments, including surgery, radiotherapy and chemotherapy, tumour recurrence is common and the prognosis is very poor. However, the timing of relapse is difficult to predict; in particular, pseudoprogression and radionecrosis after radiotherapy show enhanced lesions on MRI, which are often difficult to distinguish from glioma recurrence. The distinction between the two is of great importance for the selection of a treatment plan in clinical follow-up.2

The widely used Response Assessment in Neuro-Oncology (RANO) assessment criteria have limitations in differentiating true tumour recurrence from postradiotherapy changes. Although biopsy has a high diagnostic accuracy, as an invasive examination with many risk factors such as the need for additional surgery, sampling bias and risks of neurological complication, it is not suitable for all patients.3,4

In recent years, with the rapid development of radiological imaging technology, the potential role of PET with different tracers and advanced MRI sequences in differentiating postradiotherapy changes from true glioma...
recurrence is becoming increasingly prominent. Existing studies indicate that the combination of multiple advanced imaging methods can improve the accuracy and specificity of the diagnosis of glioma recurrence.⁵⁶ Although several studies have been published, the best imaging methods for differentiating glioma recurrence from postradiotherapy changes have not been conclusively determined. The aim of this study was to compare the diagnostic effectiveness of six imaging methods, including five conventional PET radiotracers and MRI, for differentiating glioma recurrence from postradiotherapy changes by performing a network meta-analysis (NMA). This was conducted to provide more evidence-based data for guidelines on the appropriate use of different imaging modalities for follow-up of patients with glioma after radiotherapy.

MATERIALS AND METHODS

Literature search
A systematic review and NMA were performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension statement for NMA of healthcare intervention guidelines.⁷ The original study protocol was registered with PROSPERO prior to initiation of the systematic search as an a priori study design (CRD42021293075). PubMed (August 2021), Scopus (August 2021), EMBASE (August 2021), the Web of Science (August 2021) and the Cochrane Library (August 2021) were searched independently by two investigators (with experience of >9 years and 8 years, respectively, in the field of evidence-based medicine/radiodiagnosis). The language was limited to English. The search strategy was based on the Bayes Library of Diagnostic Study and Reviews.⁸ All searches used a combination of free words and Medical Subject Headings terms. Google Scholar and the Medical Matrix search engine were used to identify relevant literature.

Selection of articles

Inclusion criteria
- Type of research: detection of glioma recurrence with different imaging modalities.
- Object of study: Patients suspected of having recurrent glioma with no age, gender, race or country restrictions.
- Reference standard for all studies: histopathological analysis (surgery, biopsy), intraoperative observation and/or follow-up examination.
- Diagnostic method: PET or PET/CT or MRI.

Exclusion criteria
- Two-by-two table of data could not be extracted.
- Abstract form only or conference proceedings.
- Single-arm study for glioma recurrence.
- Non-English articles.

Literature screening
The retrieved literature was cleared of duplicates and two reviewers independently reviewed the abstracts and further examined the full-text articles of potentially eligible citations. Any disagreement in article selection was resolved through discussion and consultation.

Data extraction
The relevant data were extracted from each study, including first author, study nation, publication year and descriptions of the study population, study design characteristics, per-lesion/per-patient studies and the reference standard. For each study, values for true positive (TP), false positive (FP), false negative (FN), true negative (TN) were extracted. If an included study had more than one observer evaluating the imaging sets, the number of TP, TN, FP and FN for each observer were calculated, and Two-by-two tables were constructed. If the studies did not report these values, two-by-two tables from the diagnostic estimates presented in the article for each index test were constructed. To resolve disagreements between reviewers, the opinion of the majority was used for further analysis.

Quality assessment
The Confidence in Network Meta-Analysis (CINeMA) tool was used to evaluate the credibility of results.⁹ CINeMA is broadly based on the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework.¹⁰ It covers six domains: (1) within-study bias (referring to the impact of risk of bias in the included studies), (2) reporting bias (referring to publication and other reporting bias), (3) indirectness, (4) imprecision, (5) heterogeneity and (6) incoherence. Reviewers assess the level of concerns for each relative effect from NMA as giving rise to ‘no concerns’, ‘some concerns’ or ‘major concerns’ in each of the six domains. Then, judgements across the domains are summarised into a single confidence rating (‘high’, ‘moderate’, ‘low’ or ‘very low’).

The six domains are relevant to all aspects of the systematic review, including literature search, data extraction and statistical analysis. Within-study bias refers to limitations that may lead to a biased estimated relative treatment effect. Reporting bias results from the inclusion of a non-representative set of studies, which may result from an incomplete literature review. Indirectness addresses the relevance of the studies, including characteristics of the population, any interventions and the outcomes of interest. A core assumption of NMA is transitivity; there is a true relative treatment effect that applies to all studies regardless of the treatments compared. Assessment of transitivity is difficult and usually involves the distribution of effect modifiers for each comparison. CINeMA’s approach to indirectness incorporates the assumption of transitivity by identifying those comparisons that may result from different definitions of the setting of interest. Assuming transitivity implies that the agreement of the estimated treatment effects is correct. This can be assessed by the incoherence domain in CINeMA.

Finally, the imprecision and heterogeneity domains address the confidence in the estimated effect and the variability in the results that contribute to each comparison¹¹;
see details in online supplemental annex 1 and online supplemental figure S1.

The Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was used to assess the risk of bias and applicability for each study.12

CINeMA tool (https://cinema.ispm.unibe.ch) and QUADAS-2 in Review Manager (V.5.3, 2014; The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark) software was used to score the included studies and assess their methodological quality. Statistical evaluation for publication bias was not performed as this is no longer recommended based on best practice recommendations by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of diagnostic test accuracy (DTA) group.13

**Data synthesis and statistical analysis**

NMA was carried out for the various imaging modalities under study. We also carried out patient-based analyses for NMA. A direct comparison was conducted using a traditional pairwise meta-analysis, and the results had a 95% credible interval in sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), accuracy and diagnostic OR (DOR); see details in online supplemental information 2. Based on the heterogeneity, the random-effect model was used for further analysis; in the absence of heterogeneity, a fixed-effect model was used. Bayesian NMA and specific graphical analysis used the ‘gemtc’ package in R software V.4.1.1 (R Foundation for Statistical Computing) and Stata V.15.1 software (StataCorp, College Station, Texas, USA).14

Using the previously described technical implementation of the Bayesian method using R software, a prior distribution (prior probability) was identified. The likelihood was calculated from the existing data and a Bayesian hierarchical model was created in NMA. Third, the prior distribution and likelihood were used in a Markov chain Monte Carlo (MCMC) simulation and a distribution that best represents the posterior distribution was set. The probability of stable distribution and the area under the probability of stable distribution and the area under the curve (SUCRA) tool was used to score the included studies and assess their methodological quality. Statistical evaluation for publication bias was not performed as this is no longer recommended based on best practice recommendations by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of diagnostic test accuracy (DTA) group.13

**RESULTS**

**Literature search and selection of studies**

A total of 32 019 potentially relevant articles were initially retrieved from PubMed, Scopus, EMBASE, the Web of Science and the Cochrane Library, and 311 studies were eligible for further review. On further inspection, 296 publications were excluded using a retrieval strategy. Ultimately, 15 studies were eligible for the final meta-analysis (figure 1); using PubMed as an example, the detailed search strategy and query terms are shown in table 1.

**Patient and public involvement**

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Figure 1** Flow chart of the study selection process.
A total of 15 articles from 9 countries with a total of 513 patients were evaluated. Four of the studies compared the diagnostic performance of 18F-FDG PET or PET/CT with MRI. Two of the studies compared the diagnostic performance of 18F-FDOPA PET/CT with MRI. Two of the studies compared the diagnostic performance of 18F-FET PET or PET/CT with MRI. Three of the studies compared the diagnostic performance of 11C-MET PET or PET/CT with MRI. Two of the studies compared the diagnostic performance of 11C-MET PET or PET/CT with MRI. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. Data extracted from these individual studies are summarised in Table 2.

Table 1 Literature search strategy

<table>
<thead>
<tr>
<th>Step no.</th>
<th>Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Glioma</td>
</tr>
<tr>
<td>#2</td>
<td>Glioma[MeSH]</td>
</tr>
<tr>
<td>#3</td>
<td>“Glial Cell Tumor”</td>
</tr>
<tr>
<td>#4</td>
<td>“Brain tumor”</td>
</tr>
<tr>
<td>#5</td>
<td>Astrocytoma</td>
</tr>
<tr>
<td>#6</td>
<td>Glioblastoma</td>
</tr>
<tr>
<td>#7</td>
<td>Recurrent</td>
</tr>
<tr>
<td>#8</td>
<td>Recurrence</td>
</tr>
<tr>
<td>#9</td>
<td>Progressive</td>
</tr>
<tr>
<td>#10</td>
<td>“After resection”</td>
</tr>
<tr>
<td>#11</td>
<td>Residual</td>
</tr>
<tr>
<td>#12</td>
<td>“Radiation injury”</td>
</tr>
<tr>
<td>#13</td>
<td>“Radiation necrosis”</td>
</tr>
<tr>
<td>#14</td>
<td>“Radiation Injuries”[MeSH]</td>
</tr>
<tr>
<td>#15</td>
<td>#1 OR #2 OR #3 OR #4 OR #5 OR #6</td>
</tr>
<tr>
<td>#16</td>
<td>#7 OR #8 OR #9 OR #10 OR #11</td>
</tr>
<tr>
<td>#17</td>
<td>#12 OR #13 OR #14</td>
</tr>
<tr>
<td>#18</td>
<td>#16 OR #17</td>
</tr>
<tr>
<td>#19</td>
<td>#15 AND #18</td>
</tr>
</tbody>
</table>

MeSH, Medical Subject Headings.

Study description
A total of 15 articles from 9 countries with a total of 513 patients were evaluated. Four of the studies compared the diagnostic performance of 18F-FDG PET or PET/CT with MRI. Two of the studies compared the diagnostic performance of 18F-FDOPA PET/CT with MRI. Two of the studies compared the diagnostic performance of 18F-FET PET or PET/CT with MRI. Three of the studies compared the diagnostic performance of 11C-MET PET or PET/CT with MRI. Two of the studies compared the diagnostic performance of 11C-MET PET or PET/CT with MRI. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. One of the studies compared the diagnostic performance of 18F-FDG with 13N-FLT. Data extracted from these individual studies are summarised in Table 2.

Certainty of the evidence assessed by GRADE approach
While in the GRADE approach appropriate accuracy studies start as high-quality evidence about diagnostic accuracy, two reviewers independently make an overall judgement on whether the quality of evidence for an outcome warrants downgrading on the basis of study limitations. The final decision showed the certainty of the evidence is moderate (figure 2).

Direct comparison
A direct pairwise comparison of the diagnostic performance of included imaging modalities for recurrent glioma showed that MRI was less sensitive and less specific than 11C-MET, 18F-FDG, 18F-FDOPA and 18F-FET.

Evidence network
Figure 3 reveals that 11C-MET, 18F-FDG PET or PET/CT and separate MRI have more studies in terms of diagnostic performances for differentiating glioma recurrence from postradiotherapy changes.

Inconsistency test
The SEN, SPE, PPV, NPV, accuracy and DOR of the included imaging modalities were analysed using inconsistency tests employing the node-splitting method, and the results indicated consistency among the direct and indirect evidence of all outcomes, and therefore, the consistency model was applied in the current study (all p>0.05).

Network meta-analysis
The trace graph shows that each MCMC chain has achieved stable fusion from the initial part, and the fluctuation of a single chain cannot be recognised by eye. The graph distribution of the density graph is basically a normal distribution, and the above results indicate that the convergence degree of the model is satisfactory. The trace and density graph of sensitivity is shown in online supplemental figures S5–S10.

As shown in figure 4, in the 95% credibility interval, the mean difference in SEN, SPE, PPV, NPV, accuracy and DOR of five different tracers PET or PET/CT imaging and MRI for differentiating glioma recurrence from postradiotherapy changes is shown. The results of the current NMA reveal that the SEN of 18F-FDG was significantly lower than that of 11C-MET, MRI, 18F-FDOPA or 18F-FET. The specificity of MRI was significantly lower than that of 18F-FDG. The NPV of 11C-MET and 18F-FDOPA was significantly higher than that of 18F-FDG. The accuracy of 18F-FDOPA and 18F-FET was significantly higher than that of 18F-FDG. There were no statistically significant differences in PPV and DOR for all imaging modalities. Overall, the combined diagnostic efficacy of 18F-FET or 18F-FDOPA PET or PET/CT imaging is superior to other methodologies.

SUCRA values
Table 3 shows the SUCRA values for five different tracers PET or PET/CT imaging and MRI for differentiating glioma recurrence from changes after radiotherapy. The results of comparisons of SUCRA values and pairwise comparisons among different diagnostic methods were consistent.

DISCUSSION
The presence of signs of recurrence in post-treatment follow-up of patients with glioma suggests a poor prognosis, with some studies showing that the median survival time for patients with first recurrence is only 9–10 months.
Table 2  Characteristics of the enrolled studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Different examination</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Nation of publication</th>
<th>Study design</th>
<th>Number of patients (M/F)</th>
<th>Average age (year)</th>
<th>Histology</th>
<th>WHO grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geetanjali Arora 2018</td>
<td>^18^F-FDG</td>
<td>24</td>
<td>2</td>
<td>5</td>
<td>8</td>
<td>India</td>
<td>Prospective</td>
<td>39 (28/11)</td>
<td>38.05</td>
<td>Astrocytoma 11, Anaplastic astrocytoma 12, Juvenile pilocytic astrocytoma 1, Diffuse fibrillary astrocytoma 1, Germiostatic astrocytoma 1, Oligodendroglioma 3, Anaplastic oligodendroglioma 3, Oligoastrocytoma 2, GBM 3, Xanthomatous astrocytoma 1, Anaplastic oligoastrocytoma 1</td>
<td>Low (I+II) (n=21), High (III+IV) (n=18)</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>25</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ryan S. Youland 2018</td>
<td>MRI</td>
<td>17</td>
<td>2</td>
<td>16</td>
<td>2</td>
<td>USA</td>
<td>Prospective</td>
<td>13 (9/4)</td>
<td>40</td>
<td>Astrocytic tumour 10, Oligoastrocytoma 3</td>
<td>Grade II (n=1), Grade III (n=5), Grade IV (n=7)</td>
</tr>
<tr>
<td></td>
<td>^18^F-DOPA</td>
<td>27</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Niels Buchmann 2018</td>
<td>^18^F-FET</td>
<td>28</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>Germany</td>
<td>Retrospective</td>
<td>32 (18/14)</td>
<td>59</td>
<td>GBM</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>14</td>
<td>0</td>
<td>18</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z.Qiao 2019</td>
<td>^11^C-MET</td>
<td>30</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>China</td>
<td>Retrospective</td>
<td>42</td>
<td>–</td>
<td>Glioma</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>22</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otakar Belohlávek 2002</td>
<td>^18^F-FDG</td>
<td>16</td>
<td>3</td>
<td>8</td>
<td>3</td>
<td>Czech Republic</td>
<td>–</td>
<td>29 (21/8)</td>
<td>–</td>
<td>Glioma</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>23</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walter Rachinger 2005</td>
<td>^18^F-FET</td>
<td>31</td>
<td>1</td>
<td>0</td>
<td>13</td>
<td>Germany</td>
<td>Retrospective</td>
<td>45 (22/23)</td>
<td>45</td>
<td>Astrocytoma 9, Oligoastrocytoma 1, Oligodendroglioma 1, Anaplastic astrocytoma 11, Oligodendroglioma 1, GBM 22</td>
<td>Grade I (n=1), Grade II (n=10), Grade III (n=12), Grade IV (n=22)</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>29</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sellam Karunanithi1 2013</td>
<td>^18^F-DOPA</td>
<td>26</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>India</td>
<td>Prospective</td>
<td>35 (28/7)</td>
<td>36.62</td>
<td>GBM 16, Astrocytoma 11, Oligodendroglioma 4, Mixed glioma 3, Others 1</td>
<td>Grade I (n=2), Grade II (n=9), Grade III (n=8), Grade IV (n=16)</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>24</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sellam Karunanithi2 2013</td>
<td>^18^F-FDG</td>
<td>10</td>
<td>0</td>
<td>11</td>
<td>7</td>
<td>India</td>
<td>Prospective</td>
<td>28 (24/4)</td>
<td>38.82</td>
<td>GBM 13, Astrocytoma 8, Oligodendroglioma 3, Mixed glioma 3, Other 1</td>
<td>Grade I (n=2), Grade II (n=13), Grade III (n=13), Grade IV (n=13)</td>
</tr>
<tr>
<td></td>
<td>^18^F-FDG</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Different examination</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Nation of publication</th>
<th>Study design</th>
<th>Number of patients (M/F)</th>
<th>Average age (year)</th>
<th>Histology</th>
<th>WHO grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amburajan Santra 2012</td>
<td>18F-FDG</td>
<td>41</td>
<td>1</td>
<td>18</td>
<td>30</td>
<td>India</td>
<td>Prospective</td>
<td>90 (66/24)</td>
<td>36.79</td>
<td>Glioma</td>
<td>Grade I (n=9)</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>56</td>
<td>24</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grade II (n=37)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grade III (n=28)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grade IV (n=16)</td>
</tr>
<tr>
<td>LI Dongli 2012</td>
<td>11C-MET</td>
<td>27</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>China</td>
<td>Retrospective</td>
<td>36</td>
<td>38</td>
<td>Glioma</td>
<td>Low (I+II) 12</td>
</tr>
<tr>
<td></td>
<td>18F-FDG</td>
<td>13</td>
<td>0</td>
<td>15</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>High (III+IV) 14</td>
</tr>
<tr>
<td>Yelda Ozsunar 2010</td>
<td>18F-FDG</td>
<td>13</td>
<td>1</td>
<td>9</td>
<td>3</td>
<td>USA</td>
<td>Retrospective</td>
<td>30</td>
<td>–</td>
<td>Glioma</td>
<td>Grade II (n=7)</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>20</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grade III (n=9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grade IV (n=19)</td>
</tr>
<tr>
<td>Michael S. Enslow 2012</td>
<td>18F-FDG</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>USA</td>
<td>–</td>
<td>15 (9/6)</td>
<td>–</td>
<td>Glioma</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>FLT</td>
<td>9</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maria M. D’Souza 2014</td>
<td>11C-MET</td>
<td>18</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>India</td>
<td>–</td>
<td>29 (20/9)</td>
<td>–</td>
<td>Anaplastic astrocytoma 16 GBM 13</td>
<td>Grades III and IV</td>
</tr>
<tr>
<td></td>
<td>MRI</td>
<td>16</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Il Ki Hong 2011</td>
<td>FLT</td>
<td>15</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>Korea</td>
<td>Retrospective</td>
<td>20 (10/10)</td>
<td>32</td>
<td>Glioma</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>18F-FDG</td>
<td>11</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.Deuschl 2017</td>
<td>MRI</td>
<td>30</td>
<td>4</td>
<td>5</td>
<td>11</td>
<td>Germany</td>
<td>Prospective</td>
<td>50 (27/23)</td>
<td>49.9</td>
<td>Glioma</td>
<td>Grade II (n=14)</td>
</tr>
<tr>
<td></td>
<td>11C-MET</td>
<td>34</td>
<td>4</td>
<td>1</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grade III (n=16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grade IV (n=20)</td>
</tr>
</tbody>
</table>

FN, false negative; FP, false positive; GBM, glioblastoma multiforme; TN, true negative; TP, true positive.
At this stage, the final diagnosis of suspected recurrence in patients with glioma still requires continuous MRI follow-up or surgical biopsy, and there is no standard imaging technique to achieve differential diagnosis. Therefore, timely selection of imaging protocols that can effectively identify glioma recurrence and postradiation changes in a timely manner can greatly reduce the hospitalisation cost, avoid psychological burden on the patient and assist in guiding clinical decisions, thus improving patient prognosis. In this paper, we compared the diagnostic efficacy of five currently used radiotracers, including $^{18}$F-FDG, $^{11}$C-MET, $^{18}$F-FDOPA, $^{18}$F-FLT and $^{18}$F-FET, in PET or PET/CT imaging and MRI in identifying glioma recurrence and postradiation changes.

From the analysis of the results of the two-pair direct pairwise comparisons and NMA comparisons, it can be concluded that the two tracers used in positron imaging, $^{18}$F-FET and $^{18}$F-FDOPA, ranked first and second in the rank probability ranking of SEN, SPE, PPV and accuracy were higher than other studies, especially $^{18}$F-FET, which ranked first with a high probability. In the area under the cumulative probability plot, the value of SUCRA is $0 \leq \text{SUCRA} \leq 1$, and when SUCRA is 1, it suggests that the diagnostic measure is absolutely valid, while when it is 0, it suggests that the diagnostic measure is absolutely invalid. $^{35}$ The ranking of diagnostic measures according to the magnitude of SUCRA values can be performed. $^{18}$F-FET ranked first in each examination in terms of SUCRA values for SEN, SPE, PPV and accuracy. $^{36}$

The results of the study show that $^{18}$F-FET PET or PET/CT had better diagnostic efficacy than the remaining four tracers and contrast-enhanced MRI, and was a more ideal imaging method to identify glioma recurrence and postradiotherapy changes; in addition, $^{18}$F-FDOPA also showed better diagnostic performance. It is worth noting that $^{11}$C-MET has a high DOR rank probability ranking and SUCRA value, and DOR is a comprehensive evaluation index that integrates SEN, SPE, PPV and NPV to indicate the chance of a positive test result as a multiple of a negative one.

Both the RANO group and the European Association for Neuro-Oncology recommend PET/CT imaging in gliomas. $^{37}$ PET/CT is a molecular imaging technique that reflects the metabolism of the lesion and has become an indispensable tool for the differential diagnosis of brain lesions in addition to MRI. $^{18}$F-FDG shows the level of tumour glucose metabolism, however, because $^{18}$F-FDG can accumulate in large amounts in normal brain tissue, the tumour-background is relatively low, and the imaging effect is poor for gliomas, especially for low-grade gliomas with low metabolic levels or lesions close to the grey matter. $^{38}$ In contrast, PET imaging with a variety of other radiotracers, such as amino acids, nucleoside analogues

---

**Figure 2** Assessment of certainty of evidence using the Grading of Recommendations, Assessment, Development and Evaluations approach.

**Figure 3** Evidence network plot of the included studies. A indicates $^{11}$C-MET PET or PET/CT; B, $^{18}$F-FDG PET or PET/CT; C, MRI; D, $^{18}$F-FDOPA PET or PET/CT; E, $^{18}$F-FET PET or PET/CT.
and spent oxygen tracers, shows new promise in the accurate identification of glioma recurrence. $^{18}$F-FET, $^{18}$F-FDOPA and $^{11}$C-MET are amino acid-based tracers that have been increasingly used for glioma imaging in recent years and are highly recommended by the RANO group. Amino acid tracers can be used in many aspects of glioma diagnosis and treatment, including tumour grading, guiding biopsy, outlining radiotherapy target areas, detecting efficacy and identifying postradiation changes and residual lesions or recurrence. The imaging principle is that the uptake of amino acid developers is relatively low in normal brain tissue, whereas the upregulation of amino acid transport proteins and increased metabolism of amino acids in tumour cells increases the uptake of radionuclide-labelled amino acid tracers by the tumour, resulting in better tumour-background contrast in PET imaging. Evangelista et al compared and concluded that $^{18}$F-FDOPA and $^{18}$F-FET/CT have similar diagnostic accuracy for high-grade glioma recurrence. Gall-diks et al reported that $^{18}$F-FET or $^{18}$F-FDOPA positron imaging studies consistently showed that both have higher diagnostic accuracy of at least 80%–90%. The results of the reticulated meta-analysis in this paper also confirms the better diagnostic efficacy of $^{18}$F-FET and $^{18}$F-FDOPA, especially $^{18}$F-FET. Ginet et al showed that for both $^{18}$F-FET and FDOPA, further analysis of the time-activity profile of tracer uptake in the tumour helped in the differential diagnosis. Pyka et al found that dynamic multiparametric analysis of $^{18}$F-FET PET can further increase its diagnostic efficacy, especially when there is a need to improve the specificity of the diagnosis. In addition, it has also been shown that $^{18}$F-FET PET is a powerful tool to discriminate glioblastoma.

<table>
<thead>
<tr>
<th>Treatments</th>
<th>SEN</th>
<th>SPE</th>
<th>PPV</th>
<th>NPV</th>
<th>Accuracy</th>
<th>DOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{11}$C-MET</td>
<td>0.69718</td>
<td>0.15435</td>
<td>0.25959</td>
<td>0.77682</td>
<td>0.59186</td>
<td>0.87721</td>
</tr>
<tr>
<td>$^{18}$F-FDG</td>
<td>0.05844</td>
<td>0.72811</td>
<td>0.67868</td>
<td>0.05999</td>
<td>0.11294</td>
<td>0.25123</td>
</tr>
<tr>
<td>MRI</td>
<td>0.38591</td>
<td>0.22906</td>
<td>0.19792</td>
<td>0.30035</td>
<td>0.31358</td>
<td>0.5384</td>
</tr>
<tr>
<td>$^{18}$F-FDOPA</td>
<td>0.81179</td>
<td>0.62498</td>
<td>0.6739</td>
<td>0.8406</td>
<td>0.83207</td>
<td>0.45442</td>
</tr>
<tr>
<td>$^{18}$F-FET</td>
<td>0.84422</td>
<td>0.73569</td>
<td>0.68052</td>
<td>0.63179</td>
<td>0.90319</td>
<td>0.48202</td>
</tr>
<tr>
<td>$^{18}$F-FLT</td>
<td>0.20248</td>
<td>0.52782</td>
<td>0.50941</td>
<td>0.39047</td>
<td>0.24637</td>
<td>0.39674</td>
</tr>
</tbody>
</table>

DOR, diagnostic OR; NPV, negative predictive value; PPV, positive predictive value; SEN, sensitivity; SPE, specificity; SUCRA, surface under the cumulative ranking curve.
recurrence from postradiation changes, and to predict overall survival cycles. 47 11C-MET has a diagnostic accuracy of around 75% for glioma recurrence 48 49 ; some studies have shown that 11C-MET has a similar high uptake of inflammation as 18F-FDG making its diagnostic specificity limited. 50 Cui et al demonstrated that PET with 18F-FET or 11C-MET had a higher sensitivity and should be combined with FDG-PET will acquire better diagnostic performance. 51

18F-FLT is an 18F-labelled nucleoside analogue that belongs to the non-amino acid PET tracers and is also often used in glioma studies with good imaging properties. 52 Thymidine is one of the nucleosides required for DNA synthesis, therefore 18F-FLT can respond to the proliferation rate of tumour cells and can be used as a marker of tumour aggressiveness. 53 18F-FLT in low-grade gliomas exhibits a correspondingly low accumulation, preventing its application for the diagnosis of recurrence in low-grade gliomas. 54 In addition, Enslow et al found no significant difference in SUVmax parameters of 18F-FLT PET/CT between glioma recurrence and post-treatment changes. 30 A meta-analysis showed that the SEN and SPE of 18F-FLT for the diagnosis of glioma recurrence was 82% and 76%, respectively, and that the overall diagnostic accuracy was improved compared with 18F-FDG. 35

MRI is now commonly used in clinical practice for routine review and follow-up of patients with glioma during and after treatment. However, the high incidence of postradiotherapy brain injury, including early pseudoprogression and late radionecrosis, and its recurrence with glioma often have similar imaging features in MRI, thus making it difficult to distinguish between the two. In recent years, the rapid development of multimodality imaging has improved the diagnostic efficacy of MRI for glioma recurrence to some extent, but its diagnostic accuracy and validity are still limited in accuracy (95% for PET, 63% for PET and 82% for MRI). 36

The GRADE approach results in an assessment of the quality of evidence in our study as moderate. We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. If the impact of this diagnostic method on patients’ important clinical outcome indicators (such as overall survival or quality of life) is further considered, this uncertainty will be even greater, so clinicians should be cautious when using this conclusion. 37 38

There are some limitations in this work; first, the number of studies with two arms and above is small, the data that can be extracted are limited and no closed loop is formed. -There is no PET/MRI literature that meets the criteria for inclusion in this study, so we hope that more relevant studies will be conducted. Second, the quality of the included literature has a GRADE rating of moderate, which may need to be confirmed by subsequent higher quality studies, and while MRI examination has integrated some new techniques, these may lead to heterogeneity within the study and a decrease in diagnostic efficacy. Finally, the methodology of diagnostic mesh meta-analysis is not well developed, and although Bayesian mesh meta-analysis is currently one of the best tools available, it may require further updates.

CONCLUSIONS

In summary, the results of the reticulated meta-analysis of this study showed that 18F-FET and 18F-FDOPA have the highest diagnostic efficacy (GRADE B) as compared with other included diagnostic methods, especially 18F-FET, and the choice should be made after comprehensive consideration in clinical practice because of the different needs for diagnostic efficacy. The above results need to be confirmed by further studies.

Contributors TX and WY performed the literature search; JL, TX and WY were involved in planning and supervised the work; JL, TX, WY, QM and XL processed the data, performed the analysis, drafted the manuscript and designed the figures. JL, TX and WY aided in interpreting the results and worked on the manuscript. JL was responsible for the conduct of the study as a guarantor. All authors discussed the results and commented on the manuscript.

Funding This work was supported by the Natural Science Foundation of Gansu Province (22JR11RA041).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID ID
Wang Yinzhong http://orcid.org/0000-0003-2208-4324

REFERENCES


3 Ellingson BM, Chung C, Pope WB, et al. Pseudoprogression, radionecrosis, inflammation or true tumor progression? challenges


