BMJ Open

Positron emission tomography (PET) for prediction of glioma histology: Protocol for an individual-level data meta-analysis of test performance.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020187
Article Type:	Protocol
Date Submitted by the Author:	01-Nov-2017
Complete List of Authors:	Trikalinos, Nikolaos; Washington University in St Louis, Internal Medicine Nihashi, Takashi ; Komaki Shimin Byoin, Department of Radiology Evangelou, Evangelos; Dept of Hygiene and Epidemiology Terasawa, Teruhiko; Fujita Hoken Eisei Daigaku Igakubu Daigakuin Igaku Kenkyuka, Department of Emergency and General Internal Medicine
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Diagnostics
Keywords:	glioma, PET, meta-analysis

Title: Positron emission tomography (PET) for prediction of glioma histology: Protocol for an individual-level data meta-analysis of test performance.

Investigators:

Nikolaos A. Trikalinos, MD[1]

Takashi Nihashi, MD, PhD[2]

Evangelos Evangelou, PhD[3,4]

Teruhiko Terasawa, MD, PhD[5]

Affiliations:

1. Department of Oncology, Washington University in St Louis School of Medicine, Missouri, USA
2. Department of Radiology, Nagoya University Graduate School of Medicine, Nagoya, Japan
3. Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece
4. Department of Epidemiology and Biostatistics, Imperial College London, London, UK

5. Section of General Internal Medicine, Department of Emergency and General Internal Medicine, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

Address correspondence to:

Nikolaos Trikalinos , MD

Division of Oncology

Washington University in St Louis

660 S. Euclid Avenue

Campus Box 8056-29

St Louis, Missouri 63110

Phone: 314.747.7955

FAX: 314.747.5123

Email: ntrikalinos@wustl.edu

Keywords: Glioma, Positron-Emission Tomography, Network Meta-Analysis

Word count: 2072 (excluding title page, abstract, references, figures and tables)

Abstract:

Introduction: Gliomas, the most commonly diagnosed primary brain tumors, are associated with varied survivals based, in part, on their histological subtype. Therefore, accurate pretreatment tumor grading is essential for patient care and clinical trial design.

Methods and Analysis: We will perform an individual-level data meta-analysis of published studies to evaluate the ability of different types of PET to differentiate high from low grade gliomas. **Ethics and Dissemination**: Ethics approval was not applicable, as this is a meta-analytic study. Results of the analysis will be submitted for publication in a peer-reviewed journal.

Review PROSPERO registration number: 78649

Article Summary:

Strengths of this study

- This is a first of its kind network meta-analysis aiming to establish the diagnostic accuracy of PET with various tracers for grading of glioma.
- Individual treatment meta-analysis with pooling of data can provide more statistical power to determine differences between imaging modalities.

Limitations

- Variability of data obtained from external sources and limited number of patients.
- Inherent limitations of MTM and IPD-MA for data interpretation / findings might be speculative

Background

Gliomas are among the most commonly diagnosed primary brain tumors with an estimated annual incidence of over 20,000 in the United States (SEER projection for 2017). Based on clinical, histological and molecular characteristics, gliomas can be broadly divided into two major clinical subcategories: low-grade and high-grade. High-grade gliomas (also called malignant gliomas) are rapidly growing tumors and include glioblastomas (grade IV), anaplastic astrocytomas (grade III), mixed anaplastic oligoastrocytomas (grade III), and anaplastic oligodendrogliomas (grade III). Despite recent advances in multimodality therapies including temozolomide, high-grade gliomas remain incurable with a median survival of less than 3 years for glioblastomas and less than 5 years for anaplastic gliomas (NCCN & ESMO Clinical Practice Guidelines). In contrast, the low-grade gliomas include astrocytomas (grade II), oligodendrogliomas (grade II), and oligoastrocytomas (grade II), and are indolent, with a median survival over 5 years[1]. Treatment of low grade gliomas is evolving and includes maximal tumor resection with the option of radiation and chemotherapy[2].

From the above it is evident that accurate pre-therapy histological grading is of paramount importance. In the usual clinical setting, histology is clarified with conventional computed tomography (CT) or magnetic resonance imaging (MRI),

image-guided biopsy /subtotal surgical resection or near total resection. Unfortunately, apart from complete resection, other approaches are error-prone: Imaging does not provide usable tissue for final diagnosis. Biopsies and partial resections can misguide, as gliomas can demonstrate histological heterogeneity due to histopathological progression; a tumor often has both low- and high-grade components. Unless total surgical resection is performed, histological results depend on which part of the tumor is sampled.

As much as it is desirable, extensive tumor resection is not always feasible in frail patients or those with tumors located adjacent to critical structures. Therefore, in order to improve tumor grading either noninvasively, or by better, targeted biopsies, one needs to be able to supplement morphological information obtained from conventional imaging with tumor-specific functional information.

PET is a useful imaging modality that provides additional metabolic information to CT- or MRI-based morphologic characterization of tumors. Its usefulness has been proven in lung cancer[3] and aggressive lymphomas[4]. Several studies have evaluated PET using various tracers such as 18F-FDG, 11C-MET, 18F-FET, and 18F-FLT for pre-therapy histological prediction to differentiate high-grade from low-grade glioma, and reported promising results[5-7]. The sample sizes of these

studies, however, are typically small, resulting in imprecise estimates of diagnostic accuracy and use heterogeneous designs / protocols for PET assessment, making interpretation of the published data difficult. Often, the positivity cutoff criteria are defined *post-hoc* to calculate the best pairs of sensitivity and specificity for each study and lack generalizability. In this situation, aggregate-level data meta-analysis (ALD-MA) typically calculates overestimated summary statistics.

A useful approach to the above problem to perform a meta-analysis based on individual-level data (Individual-level data meta-analysis; ILD-MA). This technique employs a predefined uniform cutoff value to estimate test performance measures for all included studies and combines their results. IPD-LD can also evaluate the effect of specific factors on test performance at the level of individual, simultaneously accounting for between-study variations. We therefore planned an ILD-MA to provide a comprehensive overview and quantitative synthesis of information on the test performance of PET for this purpose.

Methods

This meta-analysis is an extension of a part of a series of systematic reviews on PET in clinical management of patients with glioma[8][9]. Although these reviews share a

common literature search until June 2011, each project employs an independent prespecified research protocol with standard systematic review methodologies, and assesses mutually exclusive research objectives. Interim results of an earlier related study of this research project have been presented at an international meeting[10] but never published as full-text.

Literature search

We will use our literature database of publications on PET assessed for patients with glioma. We established the database based on the searches of PubMed and Scopus from inception through June 30, 2011 with no language restriction and full-text evaluation of potentially relevant articles found though abstract screening. The complete search strategy and full list of the database is reported elsewhere[9]. We will update the searches until July 30, 2017, and examine the reference lists of eligible studies and relevant review articles.

One reviewer (TN or NAT) will screen abstracts and at least two of three investigators (TT, NAT, and TT) examine full-text articles of potentially eligible citations. Discrepancies will be resolved by consensus.

Inclusion and exclusion criteria

We will select studies that assess PET using 18F-FDG, 11C-MET, 18F-FET, or 18F-FLT for predicting glioma grading, verified with histologic confirmation by surgery or biopsy to be eligible. We have selected these four particular tracers before conducting this research on the basis of an empirical evaluation of published studies of PET on glioma[9] and recently published narrative reviews[11],[12]. We will include both prospective and retrospective studies. We will define pathologic confirmation (either by biopsy or surgical resection) as the acceptable reference standard and explicitly exclude studies (or individual patients in a study), in which (or for whom, respectively), pathological confirmation is not performed. We will consider the total surgical resection as the (nearly) perfect reference standard, whereas, (stereotactic) biopsy will be deemed to be the imperfect reference standard. We will include only English language publications that evaluated at least 10 patients for whom PET scanning and histopathological confirmation is successfully performed. We will also exclude editorials, comments, letters to the editor, and review articles. When multiple publications with potentially overlapping patient populations are available, we will only include a publication with the largest sample size.

We will contact study authors by email if studies do not report adequate

information on PET and histological results in the participant-level. We will consider our request to be rejected if two email request reminders separately sent q 14 days after the initial contact attempt are rejected. Even in this case, we will allow for the inclusion of a study report in which quantitative data is not reported but the digital extraction (i.e., data extraction by using a digitizer from its published graphical presentations) is feasible. Otherwise we will exclude these studies. Data will be kept in a shared secure folder accessible by all co-authors.

Data extraction

One reviewer (TN or NAT) will extract descriptive data from each eligible study. Another one non-overlapping investigator (TN, NAT, or TT) will verify all the extracted data. We will extract the following published descriptive information from eligible studies: first author, year of publication, journal, patient demographics and clinical characteristics, therapeutic interventions in the case of post-therapy or recurrence assessment, technical specifications of PET, and interpretation of PET results.

One reviewer (TN, or NAT) will extract published quantitative data regarding imaging results (i.e., visual assessment and quantitative assessment such as standard uptake values (SUVs) or tumor-to-normal uptake ratio (T/N ratio)) and final diagnoses (i.e.,

histopathologic subtype and grading) at the individual level. Another one reviewer (TT) will verify all the data. We will exclude any cases in which PET scanning is unsuccessful (and thus results are arbitrary imputed *post-hoc*). We will also exclude any cases in which alternative reference standards such as clinical follow-up instead of pathological confirmation is used to determine histologic grading.

If reported, we will also extract the following individual-level variables as the candidate effect-modifiers to be evaluated in meta-regression: age, sex, clinical scenario (i.e., primary diagnosis vs. post-therapy/recurrence assessment), and tissue sampling methods (i.e, how the pathological specimens are obtained, either biopsy, partial or total resection)

Assessment of risk of bias.

To assess the risk of bias and applicability of each study, two reviewers (TN, NAT) will independently assess patient selection, index test, reference standard, and their flow and timing based on the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2). Discrepant ratings will be resolved by consensus. The complete list of operational definitions to rate each item is available from the authors upon request.

Data synthesis

We will use a bivariate model to obtain an estimate of the summary sensitivity and specificity with their corresponding confidence intervals (CIs). We will fit a two-level a generalized mixed regression model conditional on the sensitivity and specificity of each study and a bivariate normal model for the sensitivity and specificity between studies [13](Reitsma et all, 2005 Journal of Clinical Epidemiology; Chu H and Cole SR, 2006, Journal of Clinical Epidemiology). We will calculate summary positive and negative likelihood ratios based on the summary sensitivity and specificity estimates. Positive likelihood ratio (LR+) is the ratio of sensitivity over (1-specificity) whereas negative likelihood ratio (LR-) is defined as the ratio of (1-sensitivity) over specificity. The discriminating ability of a diagnostic test is better with higher LR+ and lower LR-. A good diagnostic test typically has LR+ >5.0 and LR- <0.2. We will also construct a hierarchical summary receiver operating characteristic curve (HSROC) based on the parameters of the fitted model (Rutter Cm and Gatsonis CA, 2001; Statistics in Medicine; Macaskill P, 2004; Journal of Clinical Epidemiology). We will assess between-study heterogeneity visually using forest plots, and also by plotting sensitivity and specificity in the ROC space. We will construct 95% credible regions for summary sensitivity and specificity from the estimated parameters as proposed[14].

Subgroup analysis and meta-regression

To explore heterogeneity, we will perform subgroup analyses and if feasible, the statistical differences among mutually exclusive subgroups will be assessed by univariable meta-regression. We will add a candidate modifier (as either 0 [absence] or 1 [presence]) to the test performance parameters (i.e., sensitivity and specificity parameters, or diagnostic odds ratio and threshold parameters, respectively) jointly in the bivariate random-effects model or binormal random-effects model, respectively. We will record: publication year, study design (i.e., prospective versus retrospective) as the study-level covariates, age, sex, clinical scenario (i.e., primary diagnosis vs. post-therapy/recurrence assessment) and tissue sampling methods (i.e., biopsy vs. partial or subtotal resection vs. total resection).

Comparisons among different PET methodologies

We will compare the test performance among alternative PET tracers, and also those based on different imaging assessment protocols (i.e., visual assessment vs. quantitative assessment, either T/N ratios or SUVs). Regarding the T/N ratios, we will operationally define three categories based on the specified referent tissues: mean or median uptakes

in the gray matter (GM), those in the white matter (WM), and other miscellaneous methods (MISC) including mean or median uptakes in the contra-lateral corresponding anatomical cites or those in the adjacent normal tissue regardless of GM or WM.

For indirect comparisons, we will visually compare the constructed summary ROC curves and the credible regions of sensitivity and specificity. We will also statistically assess the differences by univariable meta-regression using the tracers or assessment methods as the covariate being incorporated into the meta-analytic model. These results however, need to be interpreted carefully because findings from the indirect comparisons may be only speculative[15]. To tackle this issue, we will also perform these analyses limiting only to comparative studies that assess multiple tracers or methods for the same participants, from which direct comparisons can be performed.

We will conduct all analyses using STATA, version 14/SE (Stata Corp, College Station, TX), and OpenBUGS, version 3.2.3 (members of OpenBUGS Project Management Group; see www.openbugs.net). All tests will be two-sided and statistical significance will be defined as a p-value < .05.

Discussion

Differentiating high from low grade malignancies has significant prognostic information for a variety of tumors; gliomas are not an exception. While an optimally

resected biopsy specimen is always preferable to indirect evidence, imaging can provide helpful adjunct information. Moreover, functional imaging, such as positron emission tomography with select biotracers has the potential to become a powerful tool in assessment of patients and influence treatment decisions. By conducting a multiple treatment meta- analysis and IPD meta-analysis, this study hopes to clarify the diagnostic accuracy of PET/CT with various tracers in differentiating glioma histology and establish thresholds, upon which further studies can be performed.

Contributorship statement: Abstract screening: TN and NAT. Full-text article examination: TT, NAT, and TT. Data extraction: TN, TT and NAT. Assessment of the risk of bias and applicability of each study: TN and NAT. Statistical analysis: TT and EE. All authors (TN,TT,EE,NAT) have equally contributed to the final version of the protocol, have reviewed and have approved of it.

Competing interests: None declared.

Funding: his research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data sharing statement: Supplementary and raw data can be provided by the authors freely upon request.

Provenance and peer review: Not commissioned; externally peer reviewed.

References:

1. Claus EB, Walsh KM, Wiencke JK, Molinaro AM, Wiemels JL, Schildkraut JM, Bondy ML, Berger M, Jenkins R, Wrensch M: **Survival and low-grade glioma: the emergence of genetic information**. *Neurosurg Focus* 2015, **38**(1):E6.

2. Buckner JC, Shaw EG, Pugh SL, Chakravarti A, Gilbert MR, Barger GR, Coons S, Ricci P, Bullard D, Brown PD *et al*: **Radiation plus Procarbazine, CCNU, and Vincristine in Low-Grade Glioma**. *N Engl J Med* 2016, **374**(14):1344-1355.

3. Sheikhbahaei S, Mena E, Yanamadala A, Reddy S, Solnes LB, Wachsmann J, Subramaniam RM: **The Value of FDG PET/CT in Treatment Response Assessment, Follow-Up, and Surveillance of Lung Cancer**. *AJR Am J Roentgenol* 2017, **208**(2):420-433.

4. Adams HJ, Kwee TC: **Pretransplant FDG-PET in aggressive non-Hodgkin lymphoma: systematic review and meta-analysis**. *Eur J Haematol* 2017, **98**(4):337-347.

5. Miyake K, Shinomiya A, Okada M, Hatakeyama T, Kawai N, Tamiya T: **Usefulness of FDG, MET and FLT-PET studies for the management of human gliomas**. *J Biomed Biotechnol* 2012, **2012**:205818.

6. Sweeney R, Polat B, Samnick S, Reiners C, Flentje M, Verburg FA: **O-(2-[(18)F]fluoroethyl)-L-tyrosine uptake is an independent prognostic determinant in patients with glioma referred for radiation therapy**. *Ann Nucl Med* 2014, **28**(2):154-162.

7. Thon N, Kunz M, Lemke L, Jansen NL, Eigenbrod S, Kreth S, Lutz J, Egensperger R, Giese A, Herms J *et al*: **Dynamic 18F-FET PET in suspected WHO grade II gliomas defines distinct biological subgroups with different clinical courses**. *Int J Cancer* 2015, **136**(9):2132-2145.

8. Nihashi T, Dahabreh IJ, Terasawa T: **Diagnostic accuracy of PET for recurrent glioma diagnosis: a meta-analysis**. *AJNR Am J Neuroradiol* 2013, **34**(5):944-950, S941-911.

9. Nihashi T, Dahabreh IJ, Terasawa T: **PET in the clinical management of glioma: evidence map**. *AJR Am J Roentgenol* 2013, **200**(6):W654-660.

10. Nihashi T DI, Terasawa T.: **Positron Emission Tomography (PET) to Differentiate High-Grade from Low-Grade Glioma: A Meta-Analysis of**

Test Performance. *J Neuroimaging* 2014, **23**(2):280 (Abstract)

.

11. la Fougere C, Suchorska B, Bartenstein P, Kreth FW, Tonn JC: **Molecular imaging of gliomas with PET: opportunities and limitations**. *Neuro Oncol* 2011, **13**(8):806-819.

12. Boellaard R: **Need for standardization of 18F-FDG PET/CT for treatment response assessments**. *J Nucl Med* 2011, **52 Suppl 2**:93S-100S.

13. Riley RD, Dodd SR, Craig JV, Thompson JR, Williamson PR: **Meta-analysis of diagnostic test studies using individual patient data and aggregate data**. *Stat Med* 2008, **27**(29):6111-6136.

14. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA: **A unification of models for meta-analysis of diagnostic accuracy studies**. *Biostatistics* 2007, **8**(2):239-251.

15. Takwoingi Y, Leeflang MM, Deeks JJ: **Empirical evidence of the importance of comparative studies of diagnostic test accuracy**. *Ann Intern Med* 2013, **158**(7):544-554.

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item
ADMINISTRATIVE INFORMATION		
Title:		
Identification	1a Line1	Identify the report as a protocol of a systematic review.
Update	1b Line 123	If the protocol is for an update of a previous systematic review, identify as such
Registration	2 Line 44	If registered, provide the name of the registry (such as PROSPERO) and registration number
Authors:		
Contact	3a Line 23	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author
Contributions	3b Line 258	Describe contributions of protocol authors and identify the guarantor of the review
Amendments	4 Not applicable	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments
Support:		
Sources	5a Line 264	Indicate sources of financial or other support for the review
Sponsor	5b N/A	Provide name for the review funder and/or sponsor
Role of sponsor or funder	5c N/A	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol
INTRODUCTION		
Rationale	6 Line 72	Describe the rationale for the review in the context of what is already known
Objectives	7 Line 118	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)
METHODS		
Eligibility criteria	8 Line 132	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review
Information sources	9 Line 133	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage
Search strategy	10 Line 135	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated
Study records:		

Data management	11a Line 166	Describe the mechanism(s) that will be used to manage records and data throughout the review
Selection process	11b Line 144	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)
Data collection process	11c Line 169	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators
Data items	12 Line 171	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications
Outcomes and prioritization	13 Line 176	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale
Risk of bias in individual studies	14 Line 190	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis
Data synthesis	15a Line 205	Describe criteria under which study data will be quantitatively synthesised
	15b Line 198	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)
	15c Line 216	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned
Meta-bias(es)	16 Line 191	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)
Confidence in cumulative evidence	17 Line 211 and 217	Describe how the strength of the body of evidence will be assessed (such as GRADE)

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

Positron emission tomography (PET) for prediction of glioma histology: Protocol for an individual-level data meta-analysis of test performance.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020187.R1
Article Type:	Protocol
Date Submitted by the Author:	27-Dec-2017
Complete List of Authors:	Trikalinos, Nikolaos; Washington University in St Louis, Internal Medicine Nihashi, Takashi ; Komaki Shimin Byoin, Department of Radiology Evangelou, Evangelos; Dept of Hygiene and Epidemiology Terasawa, Teruhiko; Fujita Hoken Eisei Daigaku Igakubu Daigakuin Igaku Kenkyuka, Department of Emergency and General Internal Medicine
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Diagnostics
Keywords:	glioma, PET, meta-analysis

Title: Positron emission tomography (PET) for prediction of glioma histology: Protocol for an individual-level data meta-analysis of test performance.

Investigators:

Nikolaos A. Trikalinos, MD[1]

Takashi Nihashi, MD, PhD[2]

Evangelos Evangelou, PhD[3,4]

Teruhiko Terasawa, MD, PhD[5]

Affiliations:

1. Department of Oncology, Washington University in St Louis School of Medicine, Missouri, USA
2. Department of Radiology, Nagoya University Graduate School of Medicine, Nagoya, Japan
3. Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece
4. Department of Epidemiology and Biostatistics, Imperial College London, London, UK

5. Section of General Internal Medicine, Department of Emergency and General Internal Medicine, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

Address correspondence to:

Nikolaos Trikalinos , MD

Division of Oncology

Washington University in St Louis

660 S. Euclid Avenue

Campus Box 8056-29

St Louis, Missouri 63110

Phone: 314.747.7955

FAX: 314.747.5123

Email: ntrikalinos@wustl.edu

Keywords: Glioma, Positron-Emission Tomography, Network Meta-Analysis

Word count: 2072 (excluding title page, abstract, references, figures and tables)

Abstract:

Introduction: Gliomas, the most commonly diagnosed primary brain tumors, are associated with varied survivals based, in part, on their histological subtype. Therefore, accurate pretreatment tumor grading is essential for patient care and clinical trial design.

Methods and Analysis: We will perform an individual-level data meta-analysis of published studies to evaluate the ability of different types of PET to differentiate high from low grade gliomas. We will search PubMed and Scopus from inception through July 30, 2017 with no language restriction and full-text evaluation of potentially relevant articles. We will choose studies that assess PET using 18F-FDG, 11C-MET, 18F-FET, or 18F-FLT for grading, verified with histologic confirmation. We will include both prospective and retrospective studies. Bias will be assessed by two reviewers with the QUADAS-2 tool and as per method described by Deeks et al. **Ethics and Dissemination**: Ethics approval was not applicable, as this is a meta-analytic study. Results of the analysis will be submitted for publication in a peer-reviewed journal.

PROSPERO registration number: CRD42017078649

Article Summary:

Strengths of this study

- This is a first of its kind network meta-analysis aiming to establish the diagnostic accuracy of PET with various tracers for grading of glioma.
- Individual treatment meta-analysis with pooling of data can provide more statistical power to determine differences between imaging modalities.

Limitations

- Variability of data obtained from external sources and limited number of patients.
- Inherent limitations of MTM and IPD-MA for data interpretation / findings might be speculative

Background

Gliomas are among the most commonly diagnosed primary brain tumors with an estimated annual incidence of over 20,000 in the United States (SEER projection for 2017). Based on clinical, histological and molecular characteristics, gliomas can be broadly divided into two major clinical subcategories: low-grade and high-grade. High-grade gliomas (also called malignant gliomas) are rapidly growing tumors and include glioblastomas (grade IV), anaplastic astrocytomas (grade III), mixed anaplastic oligoastrocytomas (grade III), and anaplastic oligodendrogliomas (grade III). Despite recent advances in multimodality therapies including temozolomide, high-grade gliomas remain incurable with a median survival of less than 3 years for glioblastomas and less than 5 years for anaplastic gliomas (NCCN & ESMO Clinical Practice Guidelines). In contrast, the low-grade gliomas include astrocytomas (grade II), oligodendrogliomas (grade II), and oligoastrocytomas (grade II), and are indolent, with a median survival over 5 years[1]. Treatment of low grade gliomas is evolving and includes maximal tumor resection with the option of radiation and chemotherapy[2].

From the above it is evident that accurate pre-therapy histological grading is of paramount importance. In the usual clinical setting, histology is clarified with conventional computed tomography (CT) or magnetic resonance imaging (MRI),

image-guided biopsy /subtotal surgical resection or near total resection. Unfortunately, apart from complete resection, other approaches are error-prone: Imaging does not provide usable tissue for final diagnosis. Biopsies and partial resections can misguide, as gliomas can demonstrate histological heterogeneity due to histopathological progression; a tumor often has both low- and high-grade components. Unless total surgical resection is performed, histological results depend on which part of the tumor is sampled.

As much as it is desirable, extensive tumor resection is not always feasible in frail patients or those with tumors located adjacent to critical structures. Therefore, in order to improve tumor grading either noninvasively, or by better, targeted biopsies, one needs to be able to supplement morphological information obtained from conventional imaging with tumor-specific functional information.

PET is a useful imaging modality that provides additional metabolic information to CT- or MRI-based morphologic characterization of tumors. Its usefulness has been proven in lung cancer[3] and aggressive lymphomas[4]. Several studies have evaluated PET using various tracers such as 18F-FDG, 11C-MET, 18F-FET, and 18F-FLT for pre-therapy histological prediction to differentiate high-grade from low-grade glioma, and reported promising results[5-7]. The sample sizes of these

studies, however, are typically small, resulting in imprecise estimates of diagnostic accuracy and use heterogeneous designs / protocols for PET assessment, making interpretation of the published data difficult. Often, the positivity cutoff criteria are defined *post-hoc* to calculate the best pairs of sensitivity and specificity for each study and lack generalizability. In this situation, aggregate-level data meta-analysis (ALD-MA) typically calculates overestimated summary statistics.

A useful approach to the above problem to perform a meta-analysis based on individual-level data (Individual-level data meta-analysis; ILD-MA). This technique employs a predefined uniform cutoff value to estimate test performance measures for all included studies and combines their results. IPD-LD can also evaluate the effect of specific factors on test performance at the level of individual, simultaneously accounting for between-study variations. We therefore planned an ILD-MA to provide a comprehensive overview and quantitative synthesis of information on the test performance of PET for this purpose.

Methods

This meta-analysis is an extension of a part of a series of systematic reviews on PET in clinical management of patients with glioma[8][9]. Although these reviews share a

common literature search until June 2011, each project employs an independent prespecified research protocol with standard systematic review methodologies, and assesses mutually exclusive research objectives. Interim results of an earlier related study of this research project have been presented at an international meeting[10] but never published as full-text.

Literature search

We will use our literature database of publications on PET assessed for patients with glioma. We established the database based on the searches of PubMed and Scopus from inception through June 30, 2011 with no language restriction and full-text evaluation of potentially relevant articles found though abstract screening. The complete search strategy and full list of the database is reported elsewhere[9] and also available as a supplementary file. We will update the searches until July 30, 2017, and examine the reference lists of eligible studies and relevant review articles.

One reviewer (TN or NAT) will screen abstracts and at least two of three investigators (TT, NAT, and TT) examine full-text articles of potentially eligible citations. Discrepancies will be resolved by consensus.

Inclusion and exclusion criteria

We will select studies that assess PET using 18F-FDG, 11C-MET, 18F-FET, or 18F-FLT for predicting glioma grading, verified with histologic confirmation by surgery or biopsy to be eligible. We have selected these four particular tracers before conducting this research on the basis of an empirical evaluation of published studies of PET on glioma[9] and recently published narrative reviews[11],[12]. We will include both prospective and retrospective studies. We will define pathologic confirmation (either by biopsy or surgical resection) as the acceptable reference standard and explicitly exclude studies (or individual patients in a study), in which (or for whom, respectively), pathological confirmation is not performed. We will consider the total surgical resection as the (nearly) perfect reference standard, whereas, (stereotactic) biopsy will be deemed to be the imperfect reference standard. We will include any language publications that evaluated at least 10 patients for whom PET scanning and histopathological confirmation is successfully performed. We will exclude editorials, comments, letters to the editor, and review articles. When multiple publications with potentially overlapping patient populations are available, we will only include the publication with the largest sample size.

We will contact study authors by email if studies do not report adequate information on PET and histological results in the participant-level, or if IPD data are

not presented in the paper. We will consider our request to be rejected if two email request reminders separately sent q 14 days after the initial contact attempt are rejected. Even in this case, we will allow for the inclusion of a study report in which quantitative data is not reported but the digital extraction (i.e., data extraction by using a digitizer from its published graphical presentations) is feasible. Otherwise we will exclude these studies. Data will be kept in a shared secure folder accessible by all co-authors.

Data extraction

One reviewer (TN or NAT) will extract descriptive data from each eligible study. Another one non-overlapping investigator (TN, NAT, or TT) will verify all the extracted data. We will extract the following published descriptive information from eligible studies: first author, year of publication, journal, patient demographics and clinical characteristics, therapeutic interventions in the case of post-therapy or recurrence assessment, technical specifications of PET, and interpretation of PET results. One reviewer (TN, or NAT) will extract published quantitative data regarding imaging results (i.e., visual assessment and quantitative assessment such as standard uptake values (SUVs) or tumor-to-normal uptake ratio (T/N ratio)) and final diagnoses (i.e., histopathologic subtype and grading) at the individual level. Another one reviewer (TT) will verify all the data. We will exclude any cases in which PET scanning is

unsuccessful (and thus results are arbitrary imputed *post-hoc*). We will also exclude any cases in which alternative reference standards such as clinical follow-up instead of pathological confirmation is used to determine histologic grading.

If reported, we will also extract the following individual-level variables as the candidate effect-modifiers to be evaluated in meta-regression: age, sex, clinical scenario (i.e., primary diagnosis vs. post-therapy/recurrence assessment), and tissue sampling methods (i.e, how the pathological specimens are obtained, either biopsy, partial or total resection)

Assessment of risk of bias.

To assess the risk of bias and applicability of each study, two reviewers (TN, NAT) will independently assess patient selection, index test, reference standard, and their flow and timing based on the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2). Methods to detect publication bias are not very reliable when used in diagnostic accuracy data, especially in case of heterogeneity, but we will use the method of Deeks et al[13] that has been shown to be the least biased. Discrepant ratings will be resolved by consensus. The complete list of operational definitions to rate each item is available from the authors upon request.

Data synthesis

We will use a bivariate model to obtain an estimate of the summary sensitivity and specificity with their corresponding confidence intervals (CIs). We will fit a two-level a generalized mixed regression model conditional on the sensitivity and specificity of each study and a bivariate normal model for the sensitivity and specificity between studies [14](Reitsma et all, 2005 Journal of Clinical Epidemiology; Chu H and Cole SR, 2006, Journal of Clinical Epidemiology). We will calculate summary positive and negative likelihood ratios based on the summary sensitivity and specificity estimates. Positive likelihood ratio (LR+) is the ratio of sensitivity over (1-specificity) whereas negative likelihood ratio (LR-) is defined as the ratio of (1-sensitivity) over specificity. The discriminating ability of a diagnostic test is better with higher LR+ and lower LR-. A good diagnostic test typically has LR+ >5.0 and LR- <0.2. We will also construct a hierarchical summary receiver operating characteristic curve (HSROC) based on the parameters of the fitted model (Rutter Cm and Gatsonis CA, 2001; Statistics in Medicine; Macaskill P, 2004; Journal of Clinical Epidemiology). We will assess between-study heterogeneity visually using forest plots, and also by plotting sensitivity and specificity in the ROC space. We will construct 95% credible regions for summary

sensitivity and specificity from the estimated parameters as proposed[15].

Subgroup analysis and meta-regression

To explore heterogeneity, we will perform subgroup analyses and if feasible, the statistical differences among mutually exclusive subgroups will be assessed by univariable meta-regression. We will add a candidate modifier (as either 0 [absence] or 1 [presence]) to the test performance parameters (i.e., sensitivity and specificity parameters, or diagnostic odds ratio and threshold parameters, respectively) jointly in the bivariate random-effects model or binormal random-effects model, respectively. We will record: publication year, study design (i.e., prospective versus retrospective) as the study-level covariates, age, sex, clinical scenario (i.e., primary diagnosis vs. post-therapy/recurrence assessment) and tissue sampling methods (i.e., biopsy vs. partial or subtotal resection vs. total resection).

Comparisons among different PET methodologies

We will compare the test performance among alternative PET tracers, and also those based on different imaging assessment protocols (i.e., visual assessment vs. quantitative assessment, either T/N ratios or SUVs). Regarding the T/N ratios, we will operationally

define three categories based on the specified referent tissues: mean or median uptakes in the gray matter (GM), those in the white matter (WM), and other miscellaneous methods (MISC) including mean or median uptakes in the contra-lateral corresponding anatomical cites or those in the adjacent normal tissue regardless of GM or WM.

For indirect comparisons, we will visually compare the constructed summary ROC curves and the credible regions of sensitivity and specificity. We will also statistically assess the differences by univariable meta-regression using the tracers or assessment methods as the covariate being incorporated into the meta-analytic model. These results however, need to be interpreted carefully because findings from the indirect comparisons may be only speculative[16]. To tackle this issue, we will also perform these analyses limiting only to comparative studies that assess multiple tracers or methods for the same participants, from which direct comparisons can be performed.

We will conduct all analyses using STATA, version 14/SE (Stata Corp, College Station, TX), and OpenBUGS, version 3.2.3 (members of OpenBUGS Project Management Group; see www.openbugs.net). All tests will be two-sided and statistical significance will be defined as a p-value < .05.

Discussion

Differentiating high from low grade malignancies has significant prognostic

information for a variety of tumors; gliomas are not an exception. While an optimally resected biopsy specimen is always preferable to indirect evidence, imaging can provide helpful adjunct information. Moreover, functional imaging, such as positron emission tomography with select biotracers has the potential to become a powerful tool in assessment of patients and influence treatment decisions. By conducting a multiple treatment meta- analysis and IPD meta-analysis, this study hopes to clarify the diagnostic accuracy of PET/CT with various tracers in differentiating glioma histology and establish thresholds, upon which further studies can be performed.

Contributorship statement: Abstract screening: TN and NAT. Full-text article examination: TT, NAT, and TT. Data extraction: TN, TT and NAT. Assessment of the risk of bias and applicability of each study: TN and NAT. Statistical analysis: TT and EE. All authors (TN,TT,EE,NAT) have equally contributed to the final version of the protocol, have reviewed and have approved of it.

Competing interests: None declared.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data sharing statement: Supplementary and raw data can be provided by the authors freely upon request.

Provenance and peer review: Not commissioned; externally peer reviewed.

References:

1. Claus EB, Walsh KM, Wiencke JK, Molinaro AM, Wiemels JL, Schildkraut JM, Bondy ML, Berger M, Jenkins R, Wrensch M: **Survival and low-grade glioma: the emergence of genetic information**. *Neurosurg Focus* 2015, **38**(1):E6.
2. Buckner JC, Shaw EG, Pugh SL, Chakravarti A, Gilbert MR, Barger GR, Coons S, Ricci P, Bullard D, Brown PD *et al*: **Radiation plus Procarbazine, CCNU, and Vincristine in Low-Grade Glioma**. *N Engl J Med* 2016, **374**(14):1344-1355.
3. Sheikhbahaei S, Mena E, Yanamadala A, Reddy S, Solnes LB, Wachsmann J, Subramaniam RM: **The Value of FDG PET/CT in Treatment Response Assessment, Follow-Up, and Surveillance of Lung Cancer**. *AJR Am J Roentgenol* 2017, **208**(2):420-433.
4. Adams HJ, Kwee TC: **Pretransplant FDG-PET in aggressive non-Hodgkin lymphoma: systematic review and meta-analysis**. *Eur J Haematol* 2017, **98**(4):337-347.
5. Miyake K, Shinomiya A, Okada M, Hatakeyama T, Kawai N, Tamiya T: **Usefulness of FDG, MET and FLT-PET studies for the management of human gliomas**. *J Biomed Biotechnol* 2012, **2012**:205818.
6. Sweeney R, Polat B, Samnick S, Reiners C, Flentje M, Verburg FA: **O-(2-[(18)F]fluoroethyl)-L-tyrosine uptake is an independent prognostic determinant in patients with glioma referred for radiation therapy**. *Ann Nucl Med* 2014, **28**(2):154-162.
7. Thon N, Kunz M, Lemke L, Jansen NL, Eigenbrod S, Kreth S, Lutz J, Egensperger R, Giese A, Herms J *et al*: **Dynamic 18F-FET PET in suspected WHO grade II gliomas defines distinct biological subgroups with different clinical courses**. *Int J Cancer* 2015, **136**(9):2132-2145.
8. Nihashi T, Dahabreh IJ, Terasawa T: **Diagnostic accuracy of PET for recurrent glioma diagnosis: a meta-analysis**. *AJNR Am J Neuroradiol* 2013, **34**(5):944-950,

S941-911.

9. Nihashi T, Dahabreh IJ, Terasawa T: **PET in the clinical management of glioma: evidence map**. *AJR Am J Roentgenol* 2013, **200**(6):W654-660.

10. Nihashi T DI, Terasawa T.: **Positron Emission Tomography (PET) to Differentiate High-Grade from Low-Grade Glioma: A Meta-Analysis of Test Performance.** *J Neuroimaging* 2014, **23**(2):280 (Abstract)

.

11. la Fougere C, Suchorska B, Bartenstein P, Kreth FW, Tonn JC: **Molecular imaging of gliomas with PET: opportunities and limitations**. *Neuro Oncol* 2011, **13**(8):806-819.

12. Boellaard R: **Need for standardization of 18F-FDG PET/CT for treatment response assessments**. *J Nucl Med* 2011, **52 Suppl 2**:93S-100S.

13. Deeks JJ, Macaskill P, Irwig L: **The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed**. *J Clin Epidemiol* 2005, **58**(9):882-893.

14. Riley RD, Dodd SR, Craig JV, Thompson JR, Williamson PR: **Meta-analysis of diagnostic test studies using individual patient data and aggregate data**. *Stat Med* 2008, **27**(29):6111-6136.

15. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA: **A unification of models for meta-analysis of diagnostic accuracy studies**. *Biostatistics* 2007, **8**(2):239-251.

16. Takwoingi Y, Leeflang MM, Deeks JJ: **Empirical evidence of the importance of comparative studies of diagnostic test accuracy**. *Ann Intern Med* 2013, **158**(7):544-554.

Literature search strategies

PubMed

(("Brain Neoplasms"[Mesh] OR "Glioma"[Mesh] OR glioma* OR (glioblastoma* OR "glioblastoma multiforme") OR (astrocytoma* OR "anaplastic astrocytoma") OR (oligodendrocytoma* OR "anaplastic oligodendrocytoma") OR (brain AND (tumor* OR tumour*)) OR (neuroectodermal AND (tumor* OR tumour*)) OR ependymoma* OR oligodendroglioma*)) AND ("Tomography, Emission-Computed"[Mesh] OR (positron AND emission AND tomograph*) OR pet) AND "humans"[MeSH Terms]

SCOPUS

(TITLE-ABS-KEY(glioma*) OR TITLE-ABS-KEY(glioblastoma*) OR TITLE-ABS-KEY(glioblastoma PRE/0 multiforme) OR TITLE-ABS-KEY(astrocytoma*) OR TITLE-ABS-KEY(anaplastic PRE/0 astrocytoma) OR TITLE-ABS-KEY(oligodendrocytoma*) OR TITLE-ABS-KEY(anaplastic PRE/0 oligodendrocytoma) OR TITLE-ABS-KEY(brain tumo*r*) OR TITLE-ABS-KEY(neuroectodermal PRE/0 tumo*r*) OR TITLE-ABS-KEY(ependymoma*) OR TITLE-ABS-KEY(oligodendroglioma*)) AND (TITLE-ABS-KEY(positron PRE/0 emission PRE/0 tomograph*) OR TITLE-ABS-KEY(pet))

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item
ADMINISTRATIVE INFORMATION		
Title:		
Identification	1a Line1	Identify the report as a protocol of a systematic review.
Update	1b Line 131	If the protocol is for an update of a previous systematic review, identify as such
Registration	2 Line 49	If registered, provide the name of the registry (such as PROSPERO) and registration number
Authors:		
Contact	3a Line 23	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author
Contributions	3b Line 259	Describe contributions of protocol authors and identify the guarantor of the review
Amendments	4 Not applicable	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments
Support:		
Sources	5a Line 265	Indicate sources of financial or other support for the review
Sponsor	5b N/A	Provide name for the review funder and/or sponsor
Role of sponsor or funder	5c N/A	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol
INTRODUCTION		
Rationale	6 Line 72	Describe the rationale for the review in the context of what is already known
Objectives	7 Line 118	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)
METHODS		
Eligibility criteria	8 Line 132	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review
Information sources	9 Line 133	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage
Search strategy	10 Line 137	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated
Study records:		

Data management	11a Line 168	Describe the mechanism(s) that will be used to manage records and data throughout the review
Selection process	11b Line 143	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)
Data collection process	11c Line 168	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators
Data items	12 Line 171	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications
Outcomes and prioritization	13 Line 174	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale
Risk of bias in individual studies	14 Line 189	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis
Data synthesis	15a Line 199	Describe criteria under which study data will be quantitatively synthesised
	15b Line 204	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)
	15c Line 218	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)
	15d N/A	If quantitative synthesis is not appropriate, describe the type of summary planned
Meta-bias(es)	16 Line 194	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)
Confidence in cumulative evidence	17 Line 218	Describe how the strength of the body of evidence will be assessed (such as GRADE)

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.