McAleenan, A., Jones, H. E., Kernohan, A., Faulkner, C. L., Palmer, A., Dawson, S., Wragg, C., Jefferies, S., Brandner, S., Vale, L., Higgins, J. P. T., & Kurian, K. M. (2019). Diagnostic test accuracy and cost-effectiveness of tests for codeletion of chromosomal arms 1p and 19q in people with glioma. *Cochrane Database of Systematic Reviews*, *2019*(8), 1-25. https://doi.org/10.1002/14651858.CD013387

Cochrane Library

Cochrane Database of Systematic Reviews

Diagnostic test accuracy and cost-effectiveness of tests for codeletion of chromosomal arms 1p and 19q in people with glioma (Protocol)

McAleenan A, Jones HE, Kernohan A, Faulkner CL, Palmer A, Dawson S, Wragg C, Jefferies S, Brandner S, Vale L, Higgins JPT, Kurian KM

McAleenan A, Jones HE, Kernohan A, Faulkner CL, Palmer A, Dawson S, Wragg C, Jefferies S, Brandner S, Vale L, Higgins JPT, Kurian KM. Diagnostic test accuracy and cost-effectiveness of tests for codeletion of chromosomal arms 1p and 19q in people with glioma. *Cochrane Database of Systematic Reviews* 2019, Issue 8. Art. No.: CD013387.
DOI: 10.1002/14651858.CD013387.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Protocol]

Diagnostic test accuracy and cost-effectiveness of tests for codeletion of chromosomal arms 1p and 19q in people with glioma

Alexandra McAleenan[1], Hayley E Jones[1], Ashleigh Kernohan[2], Claire L Faulkner[3], Abigail Palmer[3], Sarah Dawson[1], Christopher Wragg[3], Sarah Jefferies[4], Sebastian Brandner[5], Luke Vale[2], Julian P T Higgins[1], Kathreena M Kurian[6]

[1]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [2]Institute of Health & Society, Newcastle University, Newcastle upon Tyne, UK. [3]Bristol Genetics Laboratory, Pathology Sciences, Southmead Hospital, Bristol, UK. [4]Department of Oncology, Addenbrooke's Hospital, Cambridge, UK. [5]Division of Neuropathology and Department of Neurodegeneration, The National Hospital for Neurology and Neurosurgery, London, UK. [6]Bristol Medical School: Brain Tumour Research Centre, Public Health Sciences, University of Bristol, Bristol, UK

Contact address: Alexandra McAleenan, Population Health Sciences, Bristol Medical School, University of Bristol, 39 Whatley Road, Bristol, BS8 2PS, UK. alexandra.mcaleenan@bristol.ac.uk.

Editorial group: Cochrane Gynaecological, Neuro-oncology and Orphan Cancer Group.
Publication status and date: New, published in Issue 8, 2019.

Citation: McAleenan A, Jones HE, Kernohan A, Faulkner CL, Palmer A, Dawson S, Wragg C, Jefferies S, Brandner S, Vale L, Higgins JPT, Kurian KM. Diagnostic test accuracy and cost-effectiveness of tests for codeletion of chromosomal arms 1p and 19q in people with glioma. *Cochrane Database of Systematic Reviews* 2019, Issue 8. Art. No.: CD013387. DOI: 10.1002/14651858.CD013387.

ABSTRACT

This is a protocol for a Cochrane Review (Diagnostic test accuracy). The objectives are as follows:

To estimate the sensitivity and specificity of each technique for determining 1p/19q codeletion status in glioma, with a view to determining the most sensitive and specific technique(s).

BACKGROUND

Gliomas are a type of brain tumour. There are different types of glioma. Gliomas may have changes in their genetic information that can help identify the specific type of glioma that is present. These changes can also give information about the likely outcomes of glioma, for example how likely it is to respond to a particular treatment. One of the possible changes that may be present is the loss of parts of chromosome 1 and chromosome 19, which is called codeletion of chromosomal arms 1p and 19q. In this review we aim to determine the most accurate way of testing whether a glioma has codeletion of chromosomal arms 1p and 19q.

Target condition being diagnosed

Gliomas are a type of brain tumour that are thought to arise from progenitor cells in the central nervous system and they share some features with glial cells. Glial cells are the cells that support and insulate neurons. Age-adjusted incidence rates for all gliomas (ICD-O-3 morphology codes 9380-9480) range from 4.67 to 5.73 per 100,000 persons (Ostrom 2014), with varied survival rates. A review of population-based studies found that pilocytic astrocytoma (World Health Organization (WHO) grade I) has the highest five-year relative survival rate, varying between 57.3% and 97.3%; whilst glioblastoma (WHO grade IV) has the poorest survival with

only 0.05% to 8.9% of people surviving five years after diagnosis (Ostrom 2014).

Complete deletion of both the short arm of chromosome 1 (1p) and the long arm of chromosome 19 (19q) (1p/19q codeletion) is a mutation that can occur in gliomas. The codeletion is thought to be an early event in the development of cancer (Pinkham 2015), and is thought to be a result of an unbalanced whole-arm translocation between chromosomes 1 and 19 with the loss of the resulting hybrid chromosome (Griffin 2006; Jenkins 2006). As described below, the codeletion has diagnostic, prognostic and predictive abilities in glioma.

According to the WHO, the diagnosis of oligodendroglioma (a subtype of glioma) and anaplastic (high-grade) oligodendroglioma requires the demonstration of both an isocitrate dehydrogenase (IDH) gene family mutation and 1p/19q codeletion (Louis 2016). A recent systematic review and meta-analysis of the prognostic value of chromosomal 1p/19q codeletion in low-grade (WHO grade II) and high-grade/anaplastic (WHO grade III) tumours found a pooled hazard ratio (HR) for mortality of 0.28 (95 % confidence Interval (CI) 0.13 to 0.62; 9 studies) favouring 1p/19q codeletion after adjusting for age, extent of resection, IDH-1 mutation and type of therapy (Hu 2016). Another systematic review and meta-analysis that evaluated the association between codeletion of 1p/19q and overall survival among people with different grades and subtypes of gliomas found that 1p/19q codeletion was associated with increased overall survival (HR 0.43; 95% CI 0.35 to 0.53; 14 studies) (Zhao 2014). Similar results were seen in both low-grade tumours (HR 0.45; 95% CI 0.30 to 0.68; 5 studies) and high-grade gliomas (HR 0.41; 95% CI 0.31 to 0.53; 6 studies). Similar results were also seen for astrocytic tumours (HR 0.52; 95% CI 0.36 to 0.75; 3 studies) and oligodendroglial tumours (HR 0.41; 95% CI 0.30 to 0.56; 9 studies) (Zhao 2014). In this review they also found that there was no difference in the HR for overall survival between studies utilising two different techniques (Polymerase Chain Reaction (PCR) and Fluorescence in situ hybridization (FISH)) to assess the status of chromosomal arms 1p and 19q (Zhao 2014).

1p/19q codeletion also predicts response to chemotherapy in anaplastic oligodendrogliomas. The European Organisation for Research and Treatment of Cancer (EORTC) study 26951 was a phase III trial comparing radiotherapy (RT) with RT plus adjuvant chemotherapy with procarbazine, lomustine and vincristine (PCV) in people with newly diagnosed anaplastic oligodendroglioma (Van Den Bent 2013). An exploratory analysis of the long-term follow-up found that there was a trend towards more benefit, in terms of increased survival, from adjuvant PCV in people with 1p/19q codeletion. In people with 1p/19q codeletion fewer than half of the people died during follow-up in the RT plus PCV group (and therefore median overall survival was not reached) versus a median survival of 112 months in the RT group (HR 0.56, 95% CI 0.31 to 1.03) (Van Den Bent 2013). In people with nondeleted 1p/19q, median overall survival was 25 months in the RT plus PCV group versus 21 months in the RT group (HR 0.83, 95% CI 0.62 to 1.10) (Van Den Bent 2013). Similarly, long-term follow-up of the Radiation Therapy Oncology Group (RTOG) study 9402, which also compared PCV plus RT with RT alone in people with pure and mixed anaplastic oligodendrogliomas, found that the median survival of those with codeleted tumours treated with PCV plus RT was twice that of people receiving RT (14.7 versus 7.3 years; HR 0.59; 95% CI, 0.37 to 0.95; P = 0 .03) (Cairncross 2013). For those with non-codeleted tumours, there was no difference in median survival by treatment arm (2.6 versus 2.7 years; HR 0.85; 95% CI, 0.58 to 1.23; P = 0.39) (Cairncross 2013).

1p/19q codeletion can be absolute, or relative if it occurs in the presence of polysomy (when cells contain at least one more copy of a chromosome than normal) or polyploidy (when cells contain more than two sets of chromosomes). Several studies have suggested that people with 1p/19q codeletions in the presence of polysomy or polyploidy have a worse prognosis (progression free survival or overall survival) than people with absolute 1p/19q codeletions, with some studies suggesting that prognosis may be similar to that of people with no codeletion (Chamberlain 2015; Jiang 2014; Ren 2013; Snuderl 2009). In all these studies, classification of polysomy occurred when more than 30% of nuclei had more than two 1q and 19p signals, as assessed by FISH. Although there are limitations to these studies, for example treatment was not standardised, these findings have led us to conclude that we are only interested in diagnosing absolute deletions.

We are also interested in diagnosing situations where one copy of 1p/19q has been lost and the other copy duplicated (also termed copy-neutral loss of heterozygosity (LOH)).

Index test(s)

This review will assess the sensitivity and specificity of any DNA-based techniques that can be used on tumour tissue to directly evaluate 1p/19q codeletion status. These include the following.

- FISH
- Chromogenic in situ hybridization (CISH)
- PCR-based LOH assays (also known as PCR-based - short tandem repeat or microsatellite analysis)
- Restriction fragment length polymorphism (RFLP) analysis
- Comparative quantitative PCR
- Multiplex-ligation-dependent probe amplification (MLPA)
- Comparative genomic hybridization (CGH)
- CGH arrays (aCGH)
- Single Nucleotide Polymorphism (SNP) arrays
- Methylation arrays
- Next-generation sequencing (NGS)

Unfortunately, there is no perfect (100% sensitive, 100% specific) 'gold standard' test for 1p/19q codeletion status: each of the above tests could theoretically produce false-positive and/or false-neg-

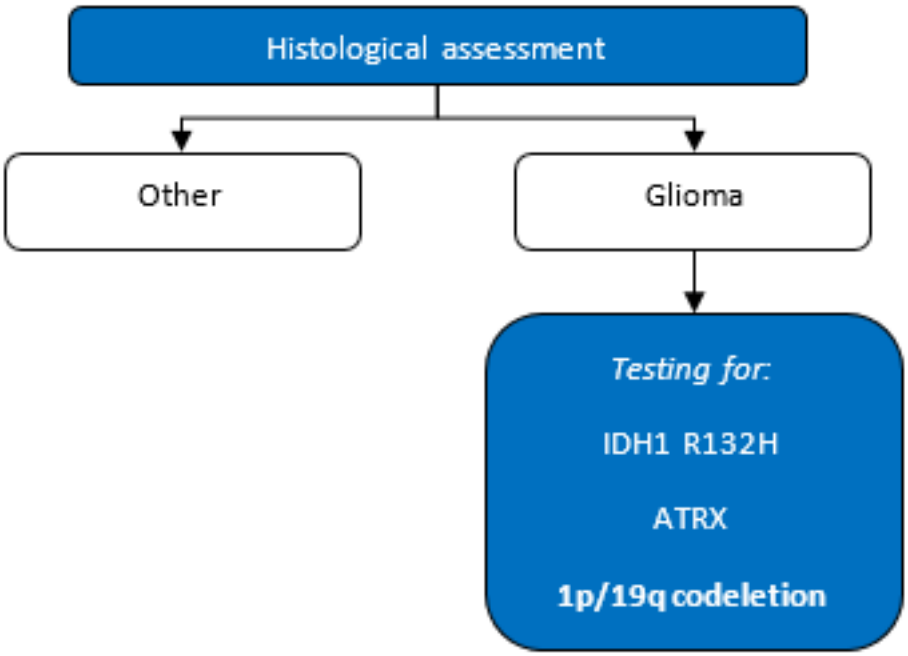
ative results. These techniques are briefly described in Table 1, along with the theoretical ways in which false-positive and false-negative results could be obtained.

Clinical pathway

Prior test(s)

Before testing for 1p/19q codeletion status, tumours undergo histological assessment. 1p/19q status is determined in tumours assessed to be gliomas by histology, as depicted in Figure 1.

Figure 1. Testing which occurs prior to determination of 1p/19q status



Role of index test(s)

As described previously, the codeletion has diagnostic, prognostic and predictive abilities in glioma.

Alternative test(s)

All DNA-based techniques which are used to determine 1p/19q status in tumour tissue will be eligible.

Rationale

European guidelines recommend that 1p/19q status is evaluated to support a diagnosis of oligodendroglioma and for prognosis, and that treatment decisions are based on the 1p/19q status (Stupp

2014; Weller 2017). WHO guidelines require the demonstration of both an IDH mutation and 1p/19q codeletion for the diagnosis of oligodendroglioma and anaplastic oligodendroglioma (Louis 2016). Recent guidance from the National Institute for Health and Care Excellence (NICE) recommend testing 1p/19q codeletion to identify oligodendrogliomas, and the adjuvant chemotherapeutic recommended after surgery for people with grade III glioma varies according to 1p/19q status (NICE 2018).

However, there are several different methods for determining 1p/19q status and no clear consensus regarding the optimal method. The two most common methods for routine diagnostic use are FISH- and PCR-based LOH assays (Woehrer 2015). In the 2017 UK Cytogenomic External Quality Assessment Service (CEQAS) report, of the 35 enrolled laboratories 25 laboratories used FISH, one laboratory used MLPA, four laboratories used arrays and one laboratory used quantitative PCR.

This review will go some way to answering the question "Do molecular subtyping techniques improve treatment selection, prediction and prognostication in people with brain and spinal cord tumours?", one of the top 10 topics identified by the James Lind Alliance Neuro-Oncology Priority Setting Partnership (James Lind Alliance). The National Cancer Research Institute Brain Clinical Studies Group has identified this as an area for future research.

A final element of the review will be to consider the costs, and cost-effectiveness of alternative methods of assessing 1p/19q status. Each method of 1p/19q assessment will incur costs, such as laboratory costs, clinic costs and subsequent treatment costs. The benefits of targeting treatment may include greater survival and less exposure to potentially toxic treatments, as well as potential cost-savings from the avoidance of waste from more efficient use of scarce health services resources.

OBJECTIVES

To estimate the sensitivity and specificity of each technique for determining 1p/19q codeletion status in glioma, with a view to determining the most sensitive and specific technique(s).

Secondary objectives

If sufficient studies are identified, we aim to break down each technique by relevant features, for example the region analysed/probes used and the cut-off used to classify 1p/19q status.

To critically appraise and summarise current evidence on the resource use, costs and cost-effectiveness of techniques for determining 1p/19q status in gliomas (conduct a full integrated review of economic evidence) and assess the cost-effectiveness of the different approaches of determining 1p/19q status using a decision model.

METHODS

Criteria for considering studies for this review

Types of studies

Types of studies for diagnostic test accuracy (DTA) review

Cross-sectional studies that use two or more tests to assess 1p/19q status in tumour tissue from the same set of people.

Types of studies for the full integrated review of economic evidence

Cost and full economic evaluations (cost-effectiveness analyses, cost-utility analyses and cost-benefit analyses) conducted alongside any study designs or as part of a modelling exercise.

Participants

Adults ($\geq$ 18 years old) with glioma.

Index tests

- Any DNA-based technique which is used to determine 1p/19q status in tumour tissue.
- Studies which assess 1p/19q status by immunohistochemistry will be excluded.
- Studies that assess 1p/19q status from blood samples or by imaging (i.e. magnetic resonance imagining, computed tomography, positron emission tomography) will be excluded.

Target conditions

Absolute 1p/19q codeletion (1p/19q codeletion in the absence of polysomy).

Reference standards

As described in Table 1, each of the tests can potentially generate false-positive and false-negative results. As such, there is no true 'gold standard' reference test. In order to estimate the sensitivity and specificity of each test, we will perform three separate sets of statistical analyses.

- Using FISH as the reference standard, i.e. assuming that FISH has 100% sensitivity and specificity
- Using PCR-based LOH assays as the reference standard, i.e. assuming that PCR-based LOH assays have 100% sensitivity and specificity

- No reference standard. Using latent class methodology, it is possible to estimate the sensitivity and specificity of a number of tests without making the strong assumption that any one test is 100% sensitive and 100% specific. Further details are provided in the 'Statistical analysis and data synthesis - DTA review' section.

Search methods for identification of studies

Electronic searches

Electronic searches for the DTA review

We will search MEDLINE Ovid (from 1946 to current date), Embase Ovid (from 1980 to current date) and BIOSIS Citation Index (from 1969 to current date). The search strategies are given in Appendix 1.

We will also search for studies available in PubMed that are not available in MEDLINE using the syntax 'pubmednotmedline[sb]'. There will be no restrictions based on language or date of publication.

Electronic searches for the full integrated review of economic evidence

Searches for economic evaluation studies will be performed in MEDLINE and Embase from April 2015. The NHS Economic evaluation database (EED) will also be searched up to the end of March 2015, when the last records were added to that database. NHS EED was based on a comprehensive search of bibliographic databases including MEDLINE and Embase so searches of MEDLINE and Embase before 1st April 2015 are not required.

Searching other resources

Searching other resources for the DTA review

We will search Open Grey (http://www.opengrey.eu/) using the free text terms from our MEDLINE search ((("chromosome 1" OR 1p) AND ("chromosome 19" OR 19q)) OR (1p19q OR "1p/19q" OR (1p* NEAR/3 19q*))) AND (glioma* OR astrocytoma* OR astroblastoma* OR ependymoma* OR subependymoma* OR oligodendroglioma* OR oligoastrocytoma* OR pleomorphic xanthoastrocytoma* OR glioblastoma* OR GBM* OR ganglioglioma* OR gliosarcoma* OR gangliocytoma* OR ((glial* OR glioneuronal* OR brain*) AND (tumor* OR tumour* OR cancer* OR neoplasm*))).

We will search for relevant material in dissertations and theses using ProQuest Dissertations & Theses Global (https://search.proquest.com/pqdtglobal/dissertations/), using the same strategy as for Open Grey but limiting to all fields except full text. We will also search the Networked Digital Library of Theses and Dissertations (http://search.ndltd.org/index.php).

The Society of Neuro-Oncology (SNO) and its partner associations the EANO and The Japan Society of Neuro-Oncology hold meetings where relevant research may be presented. We will search for abstracts from these meetings and other relevant conferences via the Web of Science Conference Proceedings Citation Index (CPCI-S) (from 1990 to current date). We will translate the BIOSIS search for CPCI-S as both databases are hosted on Web of Science.

We will also search for any ongoing studies via the WHO International Clinical Trials Registry Platform (ICTRP). The search strategy is given in Appendix 1.

We will examine the reference lists of included studies, and of any systematic reviews that we identify.

Searching other resources for the full integrated review of economic evidence

Relevant grey literature (such as health technology assessments, reports and working papers) will also be considered for inclusion.

Data collection and analysis

Selection of studies

Selection of DTA studies

Two review authors will independently screen titles and abstracts. All articles deemed to be relevant by either review author, or whose relevance cannot be determined from the abstract, will be retrieved. Full-text articles will then be independently screened in duplicate. Disagreements will be resolved by consensus, with discussion with a third review author if necessary.

Selection of economic studies

We will include full economic evaluation studies alongside any type of study design, model-based evaluations and cost-analysis will be included. For the search for economic studies, two review authors (Luke Vale (LV) and Ashleigh Kernohan (AK)) will independently screen for, identify and classify eligible studies.

The scope of this review includes multiple types of test for codeletion of chromosomal arms 1p and 19q (e.g. FISH, CISH). As a result, studies which evaluate single strategies of detection or compare multiple strategies will be included.

Data extraction and management

Two review authors will independently perform data extraction using a data extraction form. Disagreements will be resolved by consensus, with discussion with a third review author if necessary. We will extract data on the following items.

Study characteristics

- Author
- Year
- Country
- Whether the study compared two or compared multiple techniques for determining 1p/19q status

Population characteristics

- Number of participants
- Population source and setting
- Inclusion/exclusion criteria
- Tumour subtype and grade
- Prior testing
- Age
- Gender
- Karnofsky performance status
- First diagnosis or recurrent disease
- Prevalence of 1p/19 codeletion

Index tests (per test performed)

- Technique
- Tumour sample type (i.e. formalin-fixed and paraffin-embedded (FFPE) or frozen tissue)
- Region(s) analysed
- Cut-off/threshold used to determine 1p/19q status
- Method of determining threshold and whether it was pre-specified

Raw test result data

We will present the raw data from each individual study as a contingency table of cross-classified test results. For studies comparing two tests, this will be a 2 x 2 table. Regardless of whether the study treated one of the tests as a reference standard, we will not consider results to be true positives but rather 'positive on both tests'; not true negatives but rather 'negative on both tests'; not false positives but rather 'positive on test A and negative on test B'; and not false negatives but rather 'negative on test A and positive on test B'. In situations where more than two tests are compared the data will form tables of higher dimensions. For example, if three tests are compared then the table formed will be 2 x 2 x 2, i.e. eight cells of cross-classified results.

A data extraction form for economic evaluations will be developed based on the format and guidelines used to produce structured abstracts of economic evaluations for inclusion in the NHS Economic Evaluation Database (NHS EED), adapted to the specific requirements of this review. The following data will be collected from the economic evaluation studies.

- Type of evaluations
- Sources of effectiveness data
- Cost data
- Sources of cost data
- Sources of outcome valuations
- Analytical approach

Assessment of methodological quality

Assessment of methodological quality in included DTA studies

We will assess applicability and risk of bias of included studies independently and in duplicate using the QUADAS-2 tool (Whiting 2011). Disagreements will be resolved by consensus, with discussion with a third review author if necessary. We have tailored the tool to our review, and the tailored form of the tool, along with how we will judge risk of bias and applicability in each study is described in Appendix 2.

Assessment of methodological quality in included economic studies

Economic evaluation studies will be assessed for bias in two stages. The first stage will involve assessing risk of bias from the sources of the effectiveness data. If the economic evaluation is model-based then the ROBIS tool will be used to assess bias in the effectiveness studies (Whiting 2016). The second stage assessing the economic evaluations for bias is to assess the overall methodological quality of the economic component of the evaluation. Evaluations carried out alongside studies will be assessed using the CHEERS checklist (Husereau 2013) and model-based economic evaluations will be assessed using the NICE methodology checklist (NICE 2012).

Statistical analysis and data synthesis

Statistical analysis and data synthesis for the DTA review

As described above ('Reference standards' section), we will perform three separate analyses.

- Using FISH as the reference standard
- Using PCR-based LOH assays as the reference standard
- Latent class analysis

If the results obtained using the three approaches differ, we will prefer the results of the latent class analysis. This is because we do not believe the key assumption underlying analyses (1) and (2) - that either FISH or PCR-based LOH assays are 100% sensitive and 100% specific - to be plausible (see Table 1).

For analyses 1 and 2, we will perform bivariate meta-analyses of the sensitivity and specificity of each index test where possible, assuming binomial likelihoods for the number of 'true positive' and 'true negative' test results (Chu 2006; Reitsma 2005). This approach

allows for heterogeneity in sensitivity and specificity across studies and for between-study correlation in these measures. The bivariate meta-analysis model can be used to produce summary operating points (summary sensitivities and specificities) with 95% confidence, or credible regions and prediction regions. Drawing on the equivalence of the bivariate model and the hierarchical summary receiver operating characteristic (HSROC) model (Rutter 2001) in the absence of covariates, the bivariate model can also be used to produce summary receiver operating characteristic (SROC) curves (Arends 2008; Harbord 2007). For tests where commonly used thresholds are reported, we will produce summary operating points. If thresholds differ between studies, we will produce SROC curves. If a study contributes 2 x 2 contingency tables for different thresholds, we will consider using one of the recently developed extensions to the bivariate meta-analysis model that can accommodate such data (Jones 2018; Steinhauser 2016).

It is anticipated, however, that there will be few studies that compare the same index test with FISH or PCR-based LOH assays. We therefore do not expect to be able to estimate distinct between-study heterogeneity and correlation parameters for each index test. To tackle parameter identifiability, we will make assumptions where necessary such as a common amount of between-study heterogeneity in the sensitivity and specificity of all index tests, and common correlation parameters. Any such assumptions will be clearly stated and discussed. If between-study heterogeneity parameters cannot be estimated due to small numbers of studies, it may be necessary to use a fixed-effect approach. To meet the inclusion criteria, all studies must be performed in people with glioma, and therefore are likely to also be performed in tertiary care. We hope that this will reduce heterogeneity.

In addition to summary operating points, we will estimate differences in sensitivity and differences in specificity between index tests. We anticipate there will be very few direct comparative accuracy studies (studies comparing the accuracy of two index tests relative to a reference standard) (Takwoingi 2013), such that these estimated differences will mostly be informed by 'indirect' comparisons.

As described in Table 1, neither FISH nor PCR-based LOH assays are likely to be true 'gold standards': each could in principle lead to false-negative or false-positive test results. Treating these tests as gold standards could therefore lead to biased estimates of the sensitivity and specificity of all other tests. In our third set of analyses, we will therefore apply latent class meta-analysis methods (Chu 2009; Dendukuri 2012; Menten 2013; Sadatsafavi 2010; Walter 1999). These methods provide estimates of sensitivity and specificity based on a probabilistic definition of disease state. In this modelling framework, the prevalence of 1p/19q codeletion status among people with glioma in each study is treated as an additional unknown parameter to be estimated. Latent class meta-analysis methods have recently been used in a Cochrane Review that aimed to determine the accuracy of the Xpert assay for diagnosing extrapulmonary tuberculosis and rifampicin resistance (Kohli 2018).

While latent class analysis methods for a single study are well documented, application in meta-analysis is still uncommon. It has been shown that it is important to allow for tests being positively correlated within disease states, usually referred to as 'conditional dependence' (Vacek 1985). We will therefore use latent class meta-analysis methods that allow for such conditional dependencies as far as possible (Chu 2009; Dendukuri 2012; Sadatsafavi 2010). The advantage of latent class methods is that they do not make the unrealistic assumption that one of the tests is a gold standard. However, in order to relax this assumption, it is usually necessary to make some other assumptions. This is to avoid problems with parameter identifiability (Jones 2010), which are introduced by recognising that study-level prevalence and the sensitivity and specificity of the 'reference standard' are all in fact unknown.

To reduce the number of parameters that need to be estimated, we propose *a priori* that the following might be reasonable assumptions.

- That between-study heterogeneity parameters are shared across tests. Informative prior distributions may be used to constrain these heterogeneity parameters to be close to zero (where zero corresponds to a fixed-effect model).
- That the correlation between the sensitivity and specificity of any one test will be strongly negative. To aid parameter identifiability, we will consider assuming that these correlations are equal to -1.
- That people with concordant positive results with CGH, SNP arrays, methylation arrays or NGS *and* FISH (with a cut-off dependant on the percentage of cells with a 2:1 ratio) or any other technique which confirms diploidy, are almost certainly true positives. The rationale for this is that 100% specificity (no false-positive results) would be expected if a test whose only limitation is the inability to distinguish between absolute and relative deletions (for example CGH, SNP arrays, methylation arrays, NGS) were combined with a test which can detect aneuploidy, for example FISH.
- That PCR-based LOH, RFLP, SNP arrays and NGS are close to 100% sensitive (no false-negative results). False-negative results on these tests can only be obtained if there is excessive contamination of tumour samples with normal tissue. We will assume that in a research context there is likely to be great care taken to minimise the risk of contamination with normal tissue. In formative prior distributions (e.g. a uniform (0.99, 1.00) prior on sensitivity) might therefore be used to constrain the sensitivity of these tests to be at least 99% in a research context.
- The errors in tests other than FISH and CISH may be conditionally independent given the disease state. We believe it is important to account for conditional dependence in FISH and CISH results however, since these are very similar techniques.

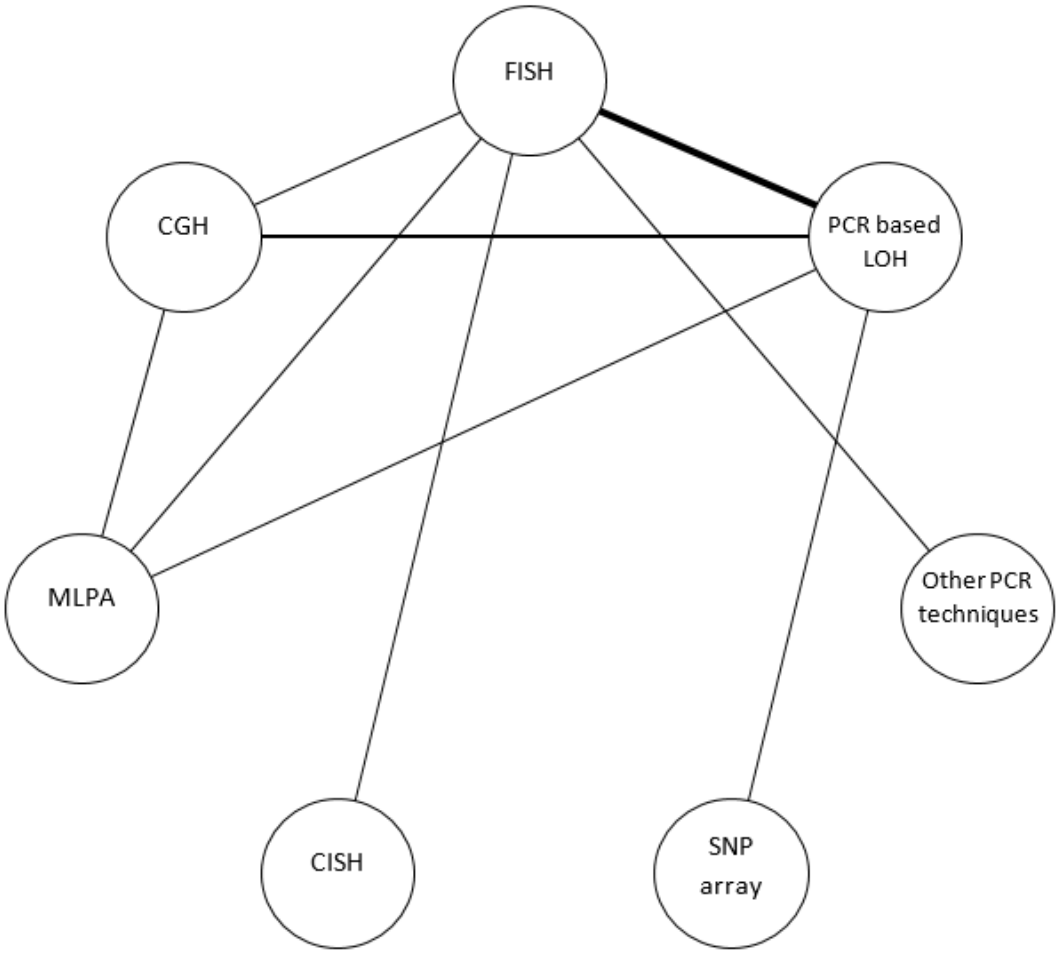
If additional tests not already described are identified that have been studied in a manner that fulfils our inclusion criteria, we will consider (i) the potential ways in which false-positive and false-

negative results could occur, in order to assess whether the test could be considered 100% sensitive or specific, and (ii) whether the test can reasonably be considered conditionally independent of other techniques given the disease state.

All statistical analyses will be carried out in the Bayesian statistical software WinBUGS (Lunn 2000). We will assume multinomial likelihoods for observed cross-classifications of test results, or binomial likelihoods for analyses assuming a perfect 'gold standard'. We will use vague prior distributions for parameters where possible. All prior distributions used will be stated, assumptions will be clearly described and model code will be provided. For the latent class analyses, we will extend WinBUGS code provided by Menten 2008 for a single study to the multiple studies setting, drawing on the methodological literature described above (Chu 2009; Dendukuri 2012; Sadatsafavi 2010).

A preliminary scoping search has identified 14 studies which have compared results from two or more different techniques (Bigner 1999; Bouvier 2004; Burger 2001; Chaturbedi 2012; Clark 2013; Harada 2011; Hatanpaa 2003; Jeuken 2006; Jha 2011; Lass 2013; Natte 2005; Nigro 2001; Scheie 2006; Smith 1999). These may be visualised as a network (see Figure 2). The only tests not included in this network are RFLP, sequencing and methylation arrays as no studies were identified in the scoping search that performed these tests and another test on the same participants.

Figure 2. Network diagram for studies identified from scoping search



The studies that we identified frequently presented results per person on each test without designating which test should be considered the reference standard. In some studies, raw results (for example PCR results from specific loci) rather than positive/negative were presented, for example in Smith 1999. In these cases we will contact authors to determine the cut-off which should be applied. In cases where no response is obtained, we will consider applying a threshold based on our own expertise, and if we are uncertain about the appropriate threshold we will exclude the data. In other studies it was clear that two different techniques had been applied to samples from the same people but usable data were not presented, for example Franco-Hernandez 2009 and Molinari 2010). Again, we will attempt to obtain raw data from the authors where possible. Similarly, if more than two tests are compared and instead of results for individual patients 2 x 2 tables for each comparison of tests are reported, we will attempt to contact authors to obtain individual patient data or fully cross-classified aggregate data. If this is unsuccessful, we will attempt to make best use of the data available.

Summary of findings for the DTA review

We will present the summary diagnostic accuracy results for each test in a 'Summary of findings' table, as illustrated in Table 2. We will prefer results obtained from latent class analyses. Confidence in each result will be assessed following the grades of recommendation, assessment, development, and evaluation (GRADE) approach (Guyatt 2008). Guidance on the use of GRADE for diagnostic tests has been published (Schunemann 2008). We will rate the overall strength of evidence as 'high', 'moderate', 'low' or 'very low' considering risk of bias, imprecision, inconsistency, applicability and publication bias, all of which may lead to downgrading the quality of the evidence. See Appendix 3.

An issue when using GRADE to rate the quality of the evidence is that test accuracy is considered a surrogate for outcomes that are important to patients and can only provide indirect evidence of impact on patient-important outcomes (Schunemann 2008). As we described previously, the codeletion has diagnostic, prognostic and predictive abilities in glioma, and all the tests described so far will have the same risk of adverse events associated with the test as they all require some biopsied tumour material. Therefore, we will assume that testing using the most accurate test will improve patient-important outcomes. We will use the indirectness domain to downgrade the quality of the evidence if studies are found to have low applicability to our review question using QUADAS-2. We will also consider publication bias, but note that there is uncertainty about the determinants of publication bias for diagnostic accuracy studies and tests for detecting funnel plot asymmetry are inadequate (Deeks 2005).

Investigations of heterogeneity

Investigations of heterogeneity planned for the DTA review

Where sufficient number of studies have assessed the same index test, we will investigate the impact of the following index test characteristics.

- Tumour sample type (i.e. FFPE or frozen tissues)
- Region(s) analysed
- Cut-off/threshold used to determine 1p/19q status

If sufficient studies are available, we will also investigate the effect of the following population characteristics.

- Study prevalence of 1p/19q codeletion
- Tumour subtype and grade

Sensitivity analyses

Sensitivity analyses planned for the DTA review

As noted above, estimates of the differences in the sensitivity and specificity of alternative tests will be informed by a mixture of direct comparative studies (where available) and indirect evidence through the network of test comparisons. Where possible, we will perform sensitivity analyses restricting only to direct comparative studies.

If sufficient data are available, we will also perform sensitivity analyses by restricting analyses to studies judged not to be at high risk of bias or low applicability.

Assessment of reporting bias

Assessment of reporting bias in the DTA review

Because of uncertainty about the determinants of publication bias for diagnostic accuracy studies and the inadequacy of tests for detecting funnel plot asymmetry (Deeks 2005), we will not perform tests aimed at detecting publication bias.

Full integrated review of economic evidence and economic model

Full integrated review of economic evidence

Characteristics and results of included economic evaluations will be summarised using additional tables, supplemented by a narrative summary that will compare and evaluate methods used and principal results between studies. This includes the currency and price year of costs, incremental cost and incremental cost-effectiveness ratios. If it is not possible to express costs in this way, then these results will be expressed as the most recent International Dollars value using implicit price deflators for GDP and GDP Purchasing Power Parities. Where possible, unit cost data will also be combined and summarised (Shemilt 2011). This review will be

conducted according to current guidance on the use of economics methods in the preparation and maintenance of Cochrane reviews (Shemilt 2011).

Economic model

The information extracted from the studies which are captured as part of this review (from both the initial and economic search) will be used to populate a decision analytic model. This model will assess the cost-effectiveness of different methods of testing for codeletion of chromosomal arms 1p and 19q in people with glioma. The outcomes will be expressed as incremental cost per true positive detected for each method. This model will include participants over the age of 18 and the model will take an NHS perspective. Parameter uncertainty will be addressed using a sensitivity analysis.

ACKNOWLEDGEMENTS

We thank Robin Grant (Co-ordinating Editor) and Gail Quinn (Managing Editor) for the Cochrane Gynaecological, Neuro-oncology and Orphan Cancer Group (GNOC) for editorial guidance. We also thank Joanne Platt (Information Specialist, GNOC) for aiding with the development of the search strategy.

We would also like to thank Hung-Yuan Cheng and Tony Ades for reading and commenting on a draft of the protocol.

We thank the Cochrane DTA team and all of our external peer reviewers, including Helen Bulbeck, Mike Jenkinson and Dwayne Boyers.

This project was supported by the National Institute for Health Research (NIHR), via Cochrane Programme Grant funding to the Cochrane Gynaecological, Neuro-oncology and Orphan Cancers Group. AM and JPTH are supported in part by Cancer Research UK (grant number C18281/A19169).The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, National Health Service (NHS) or the Department of Health, or Cancer Research UK.

REFERENCES

Additional references

Arends 2008
Arends LR, Hamza TH, van Houwelingen JC, Heijenbrok-Kal MH, Hunink MG, Stijnen T. Bivariate random effects meta-analysis of ROC curves. *Medical Decision Making* 2008;**28**(5):621–38.

Bigner 1999
Bigner SH, Matthews MR, Rasheed BK, Wiltshire RN, Friedman HS, Friedman AH, et al. Molecular genetic aspects of oligodendrogliomas including analysis by comparative genomic hybridization. *American Journal of Pathology* 1999;**155**(2):375–86.

Bouvier 2004
Bouvier C, Roll P, Quilichini B, Metellus P, Calisti A, Gilles S, et al. Deletions of chromosomes 1p and 19q are detectable on frozen smears of gliomas by FISH: usefulness for stereotactic biopsies. *Journal of Neuro-Oncology* 2004;**68**(2):141–9.

Burger 2001
Burger PC, Minn AY, Smith JS, Borell TJ, Jedlicka AE, Huntley BK, et al. Losses of chromosomal arms 1p and 19q in the diagnosis of oligodendroglioma. A study of paraffin-embedded sections. *Modern Pathology* 2001;**14**(9):842–53.

Cairncross 2013
Cairncross G, Wang M, Shaw E, Jenkins R, Brachman D, Buckner J, et al. Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: Long-term results of RTOG 9402. *Journal of Clinical Oncology* 2013;**31**(3):337–43.

Capper 2018
Capper D, Stichel D, Sahm F, Jones DT, Schrimpf D, Sill M, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. *Acta Neuropathologica* 2018;**136**(2):181–210.

Chamberlain 2015
Chamberlain MC, Born D. Prognostic significance of relative 1p/19q codeletion in oligodendroglial tumors. *Journal of Neuro-Oncology* 2015;**125**(2):249–51.

Chaturbedi 2012
Chaturbedi A, Yu L, Linskey ME, Zhou YH. Detection of 1p19q deletion by real-time comparative quantitative PCR. *Biomarker Insights* 2012;**7**:9–17.

Chu 2006
Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology* 2006;**59**(12):1331-2; author reply 1332-3.

Chu 2009
Chu HT, Chen SN, Louis TA. Random effects models in a meta-analysis of the accuracy of two diagnostic tests without a gold standard. *Journal of the American Statistical Association* 2009;**104**(486):512–23.

Clark 2013
Clark KH, Villano JL, Nikiforova MN, Hamilton RL, Horbinski C. 1p/19q testing has no significance in the workup of glioblastomas. *Neuropathology and Applied Neurobiology* 2013;**39**(6):706–17.

Deeks 2005
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. *Journal of Clinical Epidemiology* 2005;**58**(9):882–93.

Dendukuri 2012
Dendukuri N, Schiller I, Joseph L, Pai M. Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. *Biometrics* 2012; **68**(4):1285–93.

Eijk-Van Os 2011
Eijk-Van Os PGC, Schouten JP. Multiplex Ligation-dependent Probe Amplification (MLPA(R)) for the detection of copy number variation in genomic sequences. *Methods in Molecular Biology* 2011;**688**:97–126.

Franco-Hernandez 2009
Franco-Hernandez C, Martinez-Glez V, de Campos JM, Isla A, Vaquero J, Gutierrez M, et al. Allelic status of 1p and 19q in oligodendrogliomas and glioblastomas: multiplex ligation-dependent probe amplification versus loss of heterozygosity. *Cancer Genetics and Cytogenetics* 2009;**190** (2):93–6.

Griffin 2006
Griffin CA, Burger P, Morsberger L, Yonescu R, Swierczynski S, Weingart JD, et al. Identification of der(1;19)(q10; p10) in five oligodendrogliomas suggests mechanism of concurrent 1p and 19q loss. *Journal of Neuropathology and Experimental Neurology* 2006;**65**(10):988–94.

Guyatt 2008
Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**(7650):924–6.

Harada 2011
Harada S, Henderson LB, Eshleman JR, Gocke CD, Burger P, Griffin CA, et al. Genomic changes in gliomas detected using single nucleotide polymorphism array in formalin-fixed, paraffin-embedded tissue: superior results compared with microsatellite analysis. *Journal of Molecular Diagnostics* 2011;**13**(5):541–8.

Harbord 2007
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):239–51.

Hatanpaa 2003
Hatanpaa KJ, Burger PC, Eshleman JR, Murphy KM, Berg KD. Molecular diagnosis of oligodendroglioma in paraffin sections. *Laboratory Investigation* 2003;**83**(3):419–28.

Hu 2016
Hu N, Richards R, Jensen R. Role of chromosomal 1p/ 19q co-deletion on the prognosis of oligodendrogliomas: A systematic review and meta-analysis. *Interdisciplinary Neurosurgery: Advanced Techniques and Case Management* 2016;**5**:58–63.

Husereau 2013
Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* 2013;**346**:f1049.

James Lind Alliance
James Lind Alliance Priority Setting Partnerships. Neuro-oncology Top 10. http://www.jla.nihr.ac.uk/priority-setting-partnerships/neuro-oncology/top-10-priorities/ [accessed 08/10/2018].

Jenkins 2006
Jenkins RB, Blair H, Ballman KV, Giannini C, Arusell RM, Law M, et al. A t(1;19)(q10;p10) mediates the combined deletions of 1p and 19q and predicts a better prognosis of patients with oligodendroglioma. *Cancer Research* 2006;**66** (20):9852–61.

Jeuken 2006
Jeuken J, Comelissen S, Boots-Sprenger S, Gijsen S, Wesseling P. Multiplex ligation-dependent probe amplification: A diagnostic tool for simultaneous identification of different genetic markers in glial tumors. *Journal of Molecular Diagnostics* 2006;**8**(4):433–43.

Jha 2011
Jha P, Sarkar C, Pathak P, Sharma MC, Kale SS, Gupta D, et al. Detection of allelic status of 1p and 19q by microsatellite-based PCR versus FISH: Limitations and advantages in application to patient management. *Diagnostic Molecular Pathology* 2011;**20**(1):40–7.

Jiang 2014
Jiang H, Ren X, Zhang Z, Zeng W, Wang J, Lin S. Polysomy of chromosomes 1 and 19: an underestimated prognostic factor in oligodendroglial tumors. *Journal of Neuro-Oncology* 2014;**120**(1):131–8.

Jones 2010
Jones G, Johnson WO, Hanson TE, Christensen R. Identifiability of models for multiple diagnostic testing in the absence of a gold standard. *Biometrics* 2010;**66**(3): 855–63.

Jones 2018
Jones HE, Gatsonis CA, Trikalinos TA, Welton NJ, Ades AE. Quantifying how diagnostic test accuracy depends on threshold in a meta-analysis. *Diagnostic and Prognostic Research* 2018;**2**(Suppl 1):O15.

Kohli 2018
Kohli M, Schiller I, Dendukuri N, Dheda K, Denkinger CM, Schumacher SG, et al. Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. *Cochrane Database of Systematic Reviews* 2018, Issue 8. DOI: 10.1002/14651858.CD012768.pub2

Lass 2013
Lass U, Hartmann C, Capper D, Herold-Mende C, Von Deimling A, Meiboom M, et al. Chromogenic in situ hybridization is a reliable alternative to fluorescence in situ hybridization for diagnostic testing of 1p and 19q loss in paraffin-embedded gliomas. *Brain Pathology* 2013;**23**(3): 311–8.

Louis 2016
Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica* 2016; **131**(6):803–20.

Lunn 2000
Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing* 2000;**10**(4):325–37.

Menten 2008
Menten J, Boelaert M, Lesaffre E. Bayesian latent class models with conditionally dependent diagnostic tests: a case study. *Statistics in Medicine* 2008;**27**(22):4469–88.

Menten 2013
Menten J, Boelaert M, Lesaffre E. Bayesian meta-analysis of diagnostic tests allowing for imperfect reference standards. *Statistics in Medicine* 2013;**32**(30):5398–413.

Molinari 2010
Molinari C, Iorio P, Medri L, Ballardini M, Guiducci G, Cremonini AM, et al. Chromosome 1p and 19q evaluation in low-grade oligodendrogliomas: A descriptive study. *International Journal of Molecular Medicine* 2010;**25**(1): 145–51.

Natte 2005
Natte R, Van Eijk R, Eilers P, Cleton-Jansen AM, Oosting J, Kouwenhove M, et al. Multiplex ligation-dependent probe amplification for the detection of 1p and 19q chromosomal loss in oligodendroglial tumors. *Brain Pathology* 2005;**15** (3):192–7.

NICE 2012
National Institute for Health and Care Excellence. Appendices B-I. *The Guidelines Manual*. National Institute for Health and Care Excellence, 2012:1–74.

NICE 2018
National Institute for Health and Care Excellence. Brain tumours (primary) and brain metastases in adults. https: //www.nice.org.uk/guidance/ng99 [accessed 08/10/2018] 2018.

Nigro 2001
Nigro JM, Takahashi MA, Ginzinger DG, Law M, Passe S, Jenkins RB, et al. Detection of 1p and 19q loss in oligodendroglioma by quantitative microsatellite analysis, a real-time quantitative polymerase chain reaction assay. *American Journal of Pathology* 2001;**158**(4):1253–62.

Ostrom 2014
Ostrom QT, Bauchet L, Davis FG, Deltour I, Fisher JL, Langer CE, et al. The epidemiology of glioma in adults: A state of the science review. *Neuro-Oncology* 2014;**16**(7): 896–913.

Pinkham 2015
Pinkham MB, Telford N, Whitfield GA, Colaco RJ, O'Neill F, McBain CA. FISHing tips: what every clinician should know about 1p19q analysis in gliomas using Fluorescence In Situ Hybridisation. *Clinical Oncology* 2015;**27**(8):445–53.

Reitsma 2005
Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005; **58**(10):982–90.

Ren 2013
Ren X, Jiang H, Cui X, Cui Y, Ma J, Jiang Z, et al. Co-polysomy of chromosome 1q and 19p predicts worse prognosis in 1p/19q codeleted oligodendroglial tumors: FISH analysis of 148 consecutive cases. *Neuro-Oncology* 2013;**15**(9):1244–50.

Rutter 2001
Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine* 2001;**20**(19):2865–84.

Sadatsafavi 2010
Sadatsafavi M, Shahidi N, Marra F, FitzGerald MJ, Elwood KR, Guo N, et al. A statistical method was used for the meta-analysis of tests for latent TB in the absence of a gold standard, combining random-effect and latent-class methods to estimate test accuracy. *Journal of Clinical Epidemiology* 2010;**63**(3):257–69.

Scheie 2006
Scheie D, Andresen PA, Cvancarova M, Bo AS, Helseth E, Skullerud, K, et al. Fluorescence in situ hybridization (FISH) on touch preparations: A reliable method for detecting loss of heterozygosity at 1p and 19q in oligodendroglial tumors. *American Journal of Surgical Pathology* 2006;**30**(7):828–37.

Schunemann 2008
Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;**336**(7653):1106–10.

Shemilt 2011
Shemilt I, Mugford M, Byford S, Drummond M, Eisenstein E, Knapp M, et al on behalf of the Cochrane and Campbell Economics Methods Group. Chapter 15: Incorporating economics evidence. In: Higgins JPT, Green S (editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Smith 1999
Smith JS, Alderete B, Minn Y, Borell TJ, Perry A, Mohapatra G, et al. Localization of common deletion regions on 1p and 19q in human gliomas and their association with histological subtype. *Oncogene* 1999;**18**(28):4144–52.

Snuderl 2009
Snuderl M, Eichler AF, Ligon KL, Vu QU, Silver M, Betensky RA, et al. Polysomy for chromosomes 1 and 19 predicts earlier recurrence in anaplastic oligodendrogliomas with concurrent 1p/19q loss. *Clinical Cancer Research* 2009; **15**(20):6430–7.

Steinhauser 2016
Steinhauser S, Schumacher M, Rucker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. *BMC Medical Research Methodology* 2016; **16**(1):97.

Stupp 2014
Stupp R, Brada M, van den Bent MJ, Tonn JC, Pentheroudakis G, ESMO Guidelines Working Group. High-grade glioma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* 2014;**25**:93–101.

Takwoingi 2013
Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. *Annals of Internal Medicine* 2013;**158**(7):544–54.

Vacek 1985
Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics* 1985;**41**(4): 959–68.

Van Den Bent 2013
Van Den Bent MJ, Brandes AA, Taphoorn MJ, Kros JM, Kouwenhoven MC, Delattre JY, et al. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: Long-term follow-up of EORTC brain tumor group study 26951. *Journal of Clinical Oncology* 2013;**31**(3):344–50.

Walter 1999
Walter SD, Irwig L, Glasziou PP. Meta-analysis of diagnostic tests with imperfect reference standards. *Journal of Clinical Epidemiology* 1999;**52**(10):943–51.

Weller 2017
Weller M, van den Bent M, Tonn JC, Stupp R, Preusser M, Cohen-Jonathan-Moyal E, et al. European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. *Lancet Oncology* 2017;**18**(6):e315–29.

Whiting 2011
Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. Quadas-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529–36.

Whiting 2016
Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology* 2016;**69**:225–34.

Woehrer 2015
Woehrer A, Hainfellner JA. Molecular diagnostics: techniques and recommendations for 1p/19q assessment. *CNS Oncology* 2015;**4**:295–306.

Zhao 2014
Zhao J, Ma W, Zhao H. Loss of heterozygosity 1p/19q and survival in glioma: A meta-analysis. *Neuro-Oncology* 2014; **16**(1):103–12.

* *Indicates the major publication for the study*

ADDITIONAL TABLES

Table 1. Techniques that can be used to detect 1p/19q codeletion and the theoretical ways false-positive and false-negative results could be obtained.

Technique	Brief description	Potential ways false-positive results could be obtained	Potential ways false-negative results could be obtained
FISH	FISH testing uses fluorescently-labelled probes that are designed to hybridise to specific chromosomal locations. It can be performed on formalin-fixed and paraffin-embedded tissue (FFPE), and on fresh or frozen tissue. In this technique tissue architecture is preserved To test for chromosome 1p/19q codeletion chromosomes 1 and 19 are analysed on separate slides. FISH probes corresponding to regions of	Focal deletions at regions that the target probes hybridise to could lead to false-positive results as these cannot be distinguished from whole arm deletions (as only one probe per chromosome arm is normally used)	False-negative results could be obtained if there has been a loss of heterozygosity without copy number reduction

Table 1. Techniques that can be used to detect 1p/19q codeletion and the theoretical ways false-positive and false-negative results could be obtained. (*Continued*)

	1p or 19q labelled using one colour, and control probes on 1q or 19p labelled in another colour (as 1q and 19p seem to remain unaffected) are used. Many commercially available probes hybridise to loci at 1p36 and 19q13, although the FISH probes used at different centres may not target exactly the same loci (Pinkham 2015). Normal nuclei show a diploid signal ratio of 2/2 (two signals from 1p or 19q and two signals from 1q or 19p). Absolute deletions will theoretically result in 1 signal from 1p or 19q in the presence of two signals from the control loci There is no consensus on cut-offs to diagnose codeletion. This is demonstrated by the fact that the EORTC study 26951 and the RTOG study 9402 used slightly different criteria (Pinkham 2015). Some laboratories define cut-offs based on the percentage of cells with deleted and imbalanced signals, some define cut-offs based on ratios calculated by dividing the total number of test probes by the total number of control probes, and some combine percentage and ratio cut-offs		
		Depending on the way that deletions are diagnosed (i.e. the cut-off used and whether it depends on the ratio of test probes to control probes), aberrations that lead to disproportionate gain in control probe loci (i.e. 1q and 19p) could lead to false-positive results	False-negative results could be obtained if non-neoplastic nuclei are assessed
		The way that the tumour tissue is sectioned to prepare it for FISH could lead to 'truncation artefact'. Nuclei may be transected which may lead to them containing incomplete genetic material. Fals- positive results may be obtained from normal tissue in the presence of excessive truncation artefact	Excessive truncation artefact in neoplastic tissue could lead to false-negative results
CISH	This is a very similar technique to FISH, but instead of using fluorescent labelling, the probes are labelled with a marker such as biotin, digoxigenin or dinitrophenyl, and then this marker is detected using antibodies or streptavidin (that binds biotin) that is conjugated to enzymes such as horseradish peroxidase or alkaline phosphatase. The presence of the probe can then	As for FISH.	As for FISH.

Table 1. Techniques that can be used to detect 1p/19q codeletion and the theoretical ways false-positive and false-negative results could be obtained. (*Continued*)

	be visualised in the presence of a substrate that undergoes a colour change in the presence of the enzyme The advantages of CISH is that it does not require a fluorescence microscope and staining is permanent		
PCR-based LOH assays	This technique analyses polymorphic microsatellites that are dispersed throughout the genome. Different alleles have different numbers of repeats. PCR amplification of regions containing polymorphic microsatellites can therefore result in different length PCR products. If an individual is heterozygous (has two different alleles) for a microsatellite, PCR of this region will result in two different length products. If heterozygosity is lost, only one length product will be obtained. An individual must be heterozygous for a microsatellite for it to be informative, and DNA from normal tissue is required to determine this. LOH can be determined by comparing the ratio of PCR products of different lengths obtained from normal and tumour tissue Primers that amplify regions containing microsatellites on 1p and 19q can be used to determine whether 1p and 19q are codeleted. However, there is no consensus on location or number of microsatellites analysed	PCR cannot distinguish between relative and absolute deletions, so people with relative deletions will be given false-positive results	If tumour samples are heavily contaminated with normal tissue, PCR products for both alleles will be obtained in a ratio that would give a false-negative result
		Depending on primer spacing and the number of informative loci, the technique may detect focal rather than whole arm deletions	
		Imbalanced polysomy, for example gain of one copy of chromosome 1 and 19, may result in allelic imbalance and be interpreted as loss of heterozygosity	
RFLP analysis	LOH can also be detected using RFLP analysis. In RFLP restriction enzymes that recognise specific sequences are used to cut DNA, resulting in fragments of specific sizes. Different alleles may contain cut sites,	Cannot distinguish between relative and absolute deletions, so people with relative deletions will be given false-positive results	If tumour samples are heavily contaminated with normal tissue, digestion products for both alleles will be obtained in a ratio that would give a false-negative result

Table 1. Techniques that can be used to detect 1p/19q codeletion and the theoretical ways false-positive and false-negative results could be obtained. (*Continued*)

	or the DNA fragment that the restriction enzyme produces after digestion may be expected to differ due to different numbers of repeats in different alleles. Therefore, in a similar manner to PCR, LOH can be detected through loss of fragments of a specific size from informative loci (where an individual is heterozygous in normal tissue)		
		Depending on the regions analysed, it is possible that this technique may detect focal rather than whole arm deletions	
		Imbalanced polysomy, for example gain of one copy of chromosome 1 and 19, may result in allelic imbalance and be interpreted as loss of heterozygosity	
Comparative quantitative PCR	Comparative quantitative PCR compares the amount of PCR product obtained from 1p/19q with PCR product obtained from other chromosomal regions. If a deletion is present, less PCR product will be obtained. This technique has the advantages that heterozygosity at loci is not required, nor a sample of normal tissue	PCR cannot distinguish between absolute deletion and relative deletions in the presence of polyploidy (i.e. those deletions that would give a 2:4 ratio/equivalent with FISH)	If tumour samples are heavily contaminated with normal tissue the amount of PCR product obtained would result in a false-negative result
		Polysomy which causes the PCR product from control regions to increase could result in false-positive results	False-negative results could be obtained if there has been a LOH without copy number reduction
			Aneuploidy which causes the PCR product from control regions to decrease could result in false-negative results
MLPA	MLPA uses probes designed to hybridise to specific regions of the genome that have been split into two. Each probe "half" also contains sequences corresponding to universal forward and reverse binding sites for PCR primers, and one "half" contains a region of varying length to help identify the probe later The primers are hybridised to denatured sample DNA (for example from a tumour). The next step is ligation. Only probe	Cannot distinguish between absolute deletion and relative deletions in the presence of polyploidy (i.e. those deletions that would give a 2:4 ratio/equivalent with FISH)	If tumour samples are heavily contaminated with normal tissue, a false-negative result may arise

Table 1. Techniques that can be used to detect 1p/19q codeletion and the theoretical ways false-positive and false-negative results could be obtained. (*Continued*)

	halves which are hybridised to adjacent sequences on the sample DNA will be ligated together. PCR, using primers corresponding to the universal binding sides contained in the probes, is used to amplify the probes. Only those probe halves which were ligated together will be amplified to any extent, as it is only these products which contain the binding sites for both the forward and reverse PCR primers The PCR products can then be separated by length, and quantified. The results are then normalised internally (by comparing reference probes with target probes), and then compared with reference samples. Heterozygous deletions can be identified as a probe ratio of 0.5 will be observed, and heterozygous duplications from a probe ratio of 1.5. Usually, probe ratios below 0.7 or above 1.3 are regarded as indicative of a heterozygous deletion (copy number change from two to one allele) or duplication (copy number change from two to three alleles), respectively (Eijk-Van Os 2011).		
		Single nucleotide polymorphisms at primer binding sites, as single mismatches at ligation sites can inhibit ligation	False-negative results could be obtained if there has been a LOH without copy number reduction
CGH	In CGH differentially-labelled genomes from the tumour (the test genome) and normal tissue (the control genome, which does not need to be from the same person) are simultaneously hybridised to normal metaphase chromosomes. Changes in copy number, caused for example by loss or gain of regions, will alter the ratio of the two genomes. If two different fluorochromes	Cannot distinguish between absolute deletion and relative deletions in the presence of polyploidy (i.e. those deletions that would give a 2:4 ratio/ equivalent. with FISH)	If tumour samples are heavily contaminated with normal tissue, a false-negative result may arise

Table 1. Techniques that can be used to detect 1p/19q codeletion and the theoretical ways false-positive and false-negative results could be obtained. (*Continued*)

	are used to mark the genomes (or detect the labels), changes in copy number can be revealed from the relative intensities of fluorochromes used to detect the two genomes. CGH detects DNA sequence copy number changes relative to the average copy number in the entire tumour sample. However, signals can be normalised using the sex chromosomes which may help if a tumour is known to be normal for these chromosomes		False-negative results could be obtained if there has been a LOH without copy number reduction
aCGH	aCGH follow the same principles as CGH, but instead of the two genomes being competitively hybridised to metaphase chromosomes, they are hybridised to a microarray. The theoretical resolution of aCGH is greater than that of traditional CGH	As for CGH	As for CGH
SNP arrays	A SNP array is a kind of DNA microarray. SNP arrays allow both copy number status and genotype to be determined, allowing detection of both losses and copy-neutral LOH SNPs are variations at a single position in a DNA sequence. Since individuals usually inherit one copy of each SNP position from each parent, the individual's genotype at a SNP site is typically either *AA*, *AB* or *BB*. To detect abnormalities using SNP arrays, sample DNA is fragmented, labelled and hybridised to an array containing immobilised allele-specific oligonucleotide probes (one probe for each allele). The signal intensity associated with each probe is then measured. Copy number changes can be	Cannot distinguish between absolute deletion and relative deletions in the presence of polyploidy arising from whole genome duplication after the codeletion event (i.e. those deletions that would give a 2:4 ratio/equivalent with FISH)	If tumour samples are heavily contaminated with normal tissue, a false-negative result may arise

Table 1. Techniques that can be used to detect 1p/19q codeletion and the theoretical ways false-positive and false-negative results could be obtained. (*Continued*)

	detected from the intensity of signal. By comparing the result for each SNP with those from normal tissue, or by using a hidden Markov model, LOH can be detected In the rare case of 2:2 tetraploidy, it is possible that SNP arrays will not be able to distinguish absolute from relative deletions		
Methylation arrays	Genome-wide DNA methylation array data can also be used to detect 1p/19q status, as reported in Capper 2018. In methylation arrays, specific regions of the genome that may be modified by methylation are investigated. The array has two probes for each region, one for the methylated and one for unmethylated. To detect copy number variations the signal from both probes (the methylated and unmethylated) for a specific region are added together and compared with a reference genome	Cannot distinguish between absolute deletion and relative deletions in the presence of polyploidy arising from whole genome duplication after the codeletion event (i.e. those deletions that would give a 2:4 ratio/equivalent with FISH)	If tumour samples are heavily contaminated with normal tissue, a false-negative result may arise False-negative results could be obtained if there has been a LOH without copy number reduction
NGS	NGS refers to post-Sanger sequencing technologies including sequencing-by-synthesis, sequencing-by-ligation and ion semiconductor sequencing. Whilst traditional Sanger sequencing sequences a single DNA sequence, NGS is capable of sequencing multiple sequences simultaneously Techniques have been developed to detect LOH and copy number variations using NGS. Deletions can be detected by relative perturbations in the read depth LOH can be detected when the ratio of alleles at a heterozygous SNP site is perturbed	Cannot distinguish between absolute deletion and relative deletions in the presence of polyploidy arising from whole genome duplication after the codeletion event (i.e. those deletions that would give a 2:4 ratio/equivalent with FISH)	If tumour samples are heavily contaminated with normal tissue, a false-negative result may arise

aCGH: array comparative genomic hybridization; CGH: comparative genomic hybridization; CISH: chromogenic in situ hybridization; FFPE: formalin-fixed and paraffin-embedded tissue; FISH: fluorescence in situ hybridization; LOH: loss of heterozygosity; MLPA: multiplex-ligation-dependent probe amplification; NGS: next-generation sequencing; PCR: polymerase chain reaction; RLFP: restriction fragment length polymorphism; SNP: single nucleotide polymorphism.

Table 2. Proposed 'Summary of findings' table

Test	Number of participants (studies)	Mean prevalence in included studies (95% CI)	Accuracy		What do the results mean?	Quality of the evidence (GRADE)
			Sensitivity (95% CI)	Specificity (95% CI)		
					With a prevalence of X%, X people out of 100 with glioma will have a 1p/19q codeletion. Of the X people with a 1p/19q codeletion Y people will be given the correct positive result and Z people will be given a false-negative result. Of the 100-X people without the codeletion, A people will be given a correct negative result and B people will be given a false-positive result	

CI: confidence interval.

APPENDICES

Appendix 1. Database search strategies

In this review we aim to include all tests for 1p/19q codeletion which have been studied comparatively, and consequently have not got a pre-defined list of eligible index tests. We have therefore decided to focus the search strategy based on the population (people with glioma) and the target condition (codeletion of chromosomes 1 and 19).

Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily <1946 onwards>

1. exp glioma/
2. (glioma* or astrocytoma* or astroblastoma* or ependymoma* or subependymoma* or oligodendroglioma* or oligoastrocytoma* or pleomorphic xanthoastrocytoma* or glioblastoma* or GBM* or ganglioglioma* or gliosarcoma* or gangliocytoma* or ((glial* or glioneuronal* or brain*) and (tumor* or tumour* or cancer* or neoplasm*))).mp.
3. 1 or 2
4. Chromosomes, Human, Pair 1/ or (chromosome 1 or 1p).mp.
5. Chromosomes, Human, Pair 19/ or (chromosome 19 or 19q).mp.
6. (1p?19q* or "1p/19q" or (1p* adj3 19q*)).mp.
7. 4 and 5
8. 6 or 7
9. 3 and 8

Ovid Embase <1974 onwards>

1. exp glioma/
2. (glioma* or astrocytoma* or astroblastoma* or ependymoma* or subependymoma* or oligodendroglioma* or oligoastrocytoma* or pleomorphic xanthoastrocytoma* or glioblastoma* or GBM* or ganglioglioma* or gliosarcoma* or gangliocytoma* or ((glial* or glioneuronal* or brain*) and (tumor* or tumour* or cancer* or neoplasm*))).mp.
3. 1 or 2
4. Chromosome 1/ or chromosome 1p/ or (chromosome 1 or 1p).mp.
5. chromosome 19/ or chromosome 19q/ or (chromosome 19 or 19q).mp.
6. 4 and 5
7. (1p?19q* or "1p/19q" or (1p* adj3 19q*)).mp.
8. 6 or 7
9. 3 and 8

BIOSIS Citation Index <1969 onwards>

#1 TS=(glioma* or astrocytoma* or astroblastoma* or ependymoma* or subependymoma* or oligodendroglioma* or oligoastrocytoma* or pleomorphic xanthoastrocytoma* or glioblastoma* or GBM* or ganglioglioma* or gliosarcoma* or gangliocytoma* or ((glial* or glioneuronal* or brain*) and (tumor* or tumour* or cancer* or neoplasm*)))
#2 TS=(1p*19q* OR "1p/19q")
#3 TS=(("chromosome 1" OR 1p) AND ("chromosome 19" OR 19q))
#4 #2 or #3
#5 #1 and #4

WHO International Clinical Trials Registry Platform (ICTRP)

Results from each of the following search lines will be downloaded and de-duplicated in EndNote:

Search 1: *1p* and 19q**

Search 2: *1p19q or 1p/19q*

Search 3: *glioma* and diagnostic test or astrocytoma* and diagnostic test or astroblastoma* and diagnostic test or ependymoma* and diagnostic test or subependymoma* and diagnostic test or oligodendroglioma* and diagnostic test or oligoastrocytoma* and diagnostic test or pleomorphic xanthoastrocytoma* and diagnostic test or glioblastoma* and diagnostic test or GBM* and diagnostic test or ganglioglioma* and diagnostic test or gliosarcoma* and diagnostic test or gangliocytoma* and diagnostic test or glial tumor* and diagnostic test or glial tumour* and diagnostic test or glial cancer* and diagnostic test or glial neoplasm* and diagnostic test or glioneuronal tumor* and diagnostic test or glioneuronal tumour* and diagnostic test or glioneuronal cancer* and diagnostic test or glioneuronal neoplasm* and diagnostic test or brain tumor* and diagnostic test or brain tumour* and diagnostic test or brain cancer* and diagnostic test or brain neoplasm* and diagnostic test*

Search 4: *glioma* and diagnostic assessment or astrocytoma* and diagnostic assessment or astroblastoma* and diagnostic assessment or ependymoma* and diagnostic assessment or subependymoma* and diagnostic assessment or oligodendroglioma* and diagnostic assessment or oligoastrocytoma* and diagnostic assessment or pleomorphic xanthoastrocytoma* and diagnostic assessment or glioblastoma* and diagnostic*

assessment or GBM and diagnostic assessment or ganglioglioma* and diagnostic assessment or gliosarcoma* and diagnostic assessment or gangliocytoma* and diagnostic assessment or glial tumor* and diagnostic assessment or glial tumour* and diagnostic assessment or glial cancer* and diagnostic assessment or glial neoplasm* and diagnostic assessment or glioneuronal tumor* and diagnostic assessment or glioneuronal tumour* and diagnostic assessment or glioneuronal cancer* and diagnostic assessment or glioneuronal neoplasm* and diagnostic assessment or brain tumor* and diagnostic assessment or brain tumour* and diagnostic assessment or brain cancer* and diagnostic assessment or brain neoplasm* and diagnostic assessment or brainstem tumor* and diagnostic assessment or brainstem tumour* and diagnostic assessment or brainstem cancer* and diagnostic assessment or brainstem neoplasm* and diagnostic assessment*

Search 5: *glioma* and DTA or astrocytoma* and DTA or astroblastoma* and DTA or ependymoma* and DTA or subependymoma* and DTA or oligodendroglioma* and DTA or oligoastrocytoma* and DTA or pleomorphic xanthoastrocytoma* and DTA or glioblastoma* and DTA or GBM* and DTA or ganglioglioma* and DTA or gliosarcoma* and DTA or gangliocytoma* and DTA or glial tumor* and DTA or glial tumour* and DTA or glial cancer* and DTA or glial neoplasm* and DTA or glioneuronal tumor* and DTA or glioneuronal tumour* and DTA or glioneuronal cancer* and DTA or glioneuronal neoplasm* and DTA or brain tumor* and DTA or brain tumour* and DTA or brain cancer* and DTA or brain neoplasm* and DTA or brainstem tumor* and DTA or brainstem tumour* and DTA or brainstem cancer* and DTA or brainstem neoplasm* and DTA*

Appendix 2. Review-specific tailoring of QUADAS-2

Domain 1: Patient selection

Risk of bias

Was a consecutive or random sample of patients enrolled?

Yes: If a consecutive sample or a random sample of eligible participants was included in the study

No: If a non-consecutive sample or a non-random sample of eligible participants was included in the study

Unclear: If it is not clear whether a consecutive sample or a random sample of eligible participants was included in the study

Was a case-control (or 'two-gate') design avoided?

Yes: If the study had a single set of inclusion criteria

No: If the study had more than one set of inclusion criteria

Unclear: If the inclusion criteria for the study are not clear

Did the study avoid inappropriate exclusions?

Yes: If all patients with glioma were included

No: If a subset of patients with glioma were excluded due to sub-classification/severity of glioma

Unclear: If the inclusion criteria for the study are not clear

Overall: Could the selection of patients have introduced bias?

We will take highest concern from any individual signalling question as our overall judgement (i.e. risk of bias will be classified as low if the response to all three questions is 'yes'; high if the response to any question is 'no'; and unclear if the response to any question is 'unclear' and the criteria for high risk of bias are not fulfilled).

Applicability

Are there concerns that the included patients do not match the review question?

High: If the study population included patients who would not have undergone testing in real practice, for example healthy controls.

Low: If the study included only a clinically relevant population that would have undergone testing in real practice

Unclear: If the inclusion criteria for the study are not clear

Domain 2: Index test

Risk of bias

Were the index test results interpreted without knowledge of the results of the other tests being compared?

Yes: If the index test is objective or if subjective if interpreted without the knowledge of the results of other tests for 1p/19q codeletion. The first test to be interpreted will be judged to be interpreted without knowledge of the results of the other tests even if it is not explicitly reported that it was interpreted 'blind' or without the knowledge of other test results

No: If test is subjective and interpreted with the knowledge of the results of other tests for 1p/19q codeletion

Unclear: If the test is subjective and it was unclear whether it was interpreted with the knowledge of other tests for 1p/19q codeletion

If a threshold was used, was it prespecified?

Yes: If the definition of what was considered to be a positive test result was defined before testing was performed

No: If the definition of what was considered to be a positive test result was defined after testing was performed and based on the results

Unclear: If it was unclear whether the definition of what was considered to be a positive test result was defined before testing was performed

Overall: Could the conduct or interpretation of the index test have introduced bias?

We will take highest concern from any individual signalling question as our overall judgement.

Applicability: Are there concerns that the index test, its conduct, or its interpretation differ from the review question?

High: If there are concerns that the index test, its conduct, or its interpretation differ from the review question

Low: If there are no concerns that the index test, its conduct, or its interpretation differ from the review question

Unclear: If the description of the index test is inadequate

Domain 3: Reference standard

We envisage that many studies will have compared two or more tests without necessarily designating a reference standard.

In addition, as we are planning to perform a latent class analysis, which allows for an imperfect reference standard, the risk of bias signalling question regarding whether the reference standard is likely to correctly classify the target condition can be omitted.

We have also decided that the applicability question is not relevant.

We have therefore decided to complete domain 2 for each test that has been compared.

Domain 4: Flow and timing

We have modified some of the wording of the signalling questions to reflect the fact that studies may not have designated a reference standard, and that this is not an issue for our latent class analysis.

Risk of bias

Was there an appropriate interval between the tests being compared?

We envisage that most tests will be done on biopsied material.

Yes: If all tests were performed on biopsied tumour material collected on one occasion

No: If test were performed on tumour material collected at different time points

Unclear: If it is unclear whether the tests were performed on the same material.

Were all patients included in the analysis?

Yes: If all participants were included in the analysis, or if participants were excluded because they did not meet inclusion criteria or if withdrawals were less than 5% of the enrolled population (arbitrarily selected cut-off)

No: If any participants were excluded from the analysis because of uninterpretable results, because of inability to undergo index test or reference standard or for unclear reasons

Overall: Could the patient flow have introduced bias?

We will take highest concern from any individual signalling question as our overall judgement.

Appendix 3. Domains to be considered when judging the strength of the body of evidence

Domains to be considered when judging the strength of the body of evidence, based on GRADE.

Domain	Explanation
Risk of bias	Based on results of risk of bias assessments. Confidence in the evidence base will be downgraded if most of the evidence is from studies not judged to be at low risk of bias
Imprecision	Confidence in the evidence base will be downgraded if the estimate of the effect size from a meta-analysis is not precise, or, if no meta-analysis is performed, if the estimate of the size of effect from individual studies is not precise
Inconsistency	Confidence in the evidence base will be downgraded if there is unexplained heterogeneity or variability in results across studies
Indirectness	Based on QUADAS-2 assessments of applicability. Confidence in the evidence base will be downgraded if most of the evidence is from studies judged to have low applicability to the review question

(*Continued*)

Publication bias	Confidence in the evidence base will be downgraded if we uncover evidence of publication bias

CONTRIBUTIONS OF AUTHORS

AM and HEJ drafted the diagnostic test accuracy sections of the protocol. JPTH provided methodological guidance. KMK, SB, SJ provided clinical guidance. SD provided searching guidance. AK and LV drafted the economic sections of the protocol. All review authors commented on and approved the protocol.

DECLARATIONS OF INTEREST

Alexandra McAleenan: none known

Hayley E Jones: none known

Ashleigh Kernohan: none known

Claire L Faulkner: none known

Abigail Palmer: none known

Sarah Dawson: none known

Christopher Wragg: none known

Sarah Jefferies: none known

Sebastian Brandner: none known

Luke Vale: none known

Julian PT Higgins: none known

Kathreena M Kurian: none known

SOURCES OF SUPPORT

Internal sources

- University of Bristol, UK.

External sources

- National Institute for Health Research, UK.

This work is supported by a National Institute for Health Research Systematic Reviews Cochrane Programme Grant 16/144

- Cancer Research UK, UK.

AM and JPTH are supported in part by Cancer Research UK (grant number C18281/A19169)

The views and opinions expressed are those of the review authors and do not necessarily reflect those of the NIHR, its Systematic Reviews Programme, the National Health Service (NHS), the Department of Health, or Cancer Research UK.