![Cochrane Library logo]
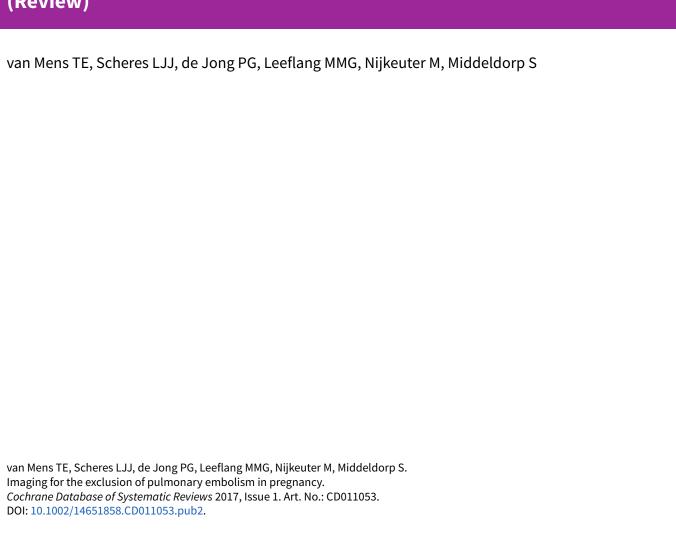
Cochrane Database of Systematic Reviews

Imaging for the exclusion of pulmonary embolism in pregnancy (Review)

van Mens TE, Scheres LJJ, de Jong PG, Leeflang MMG, Nijkeuter M, Middeldorp S

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Imaging for the exclusion of pulmonary embolism in pregnancy

Thijs E van Mens[1], Luuk JJ Scheres[1], Paulien G de Jong[1], Mariska MG Leeflang[2], Mathilde Nijkeuter[1,3], Saskia Middeldorp[1]

[1]Department of Vascular Medicine, Academic Medical Center, Amsterdam, Netherlands. [2]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands. [3]Department of Internal Medicine, University Medical Centre Utrecht, Utrecht, Netherlands

Contact address: Thijs E van Mens, Department of Vascular Medicine, Academic Medical Center, Meibergdreef 9, Amsterdam, 1105 AZ, Netherlands. t.e.vanmens@amc.nl, thijsvanmens@hotmail.com.

A B S T R A C T

Background

Pulmonary embolism is a leading cause of pregnancy-related death. An accurate diagnosis in pregnant patients is crucial to prevent untreated pulmonary embolism as well as unnecessary anticoagulant treatment and future preventive measures. Applied imaging techniques might perform differently in these younger patients with less comorbidity and altered physiology, who largely have been excluded from diagnostic studies.

Objectives

To determine the diagnostic accuracy of computed tomography pulmonary angiography (CTPA), lung scintigraphy and magnetic resonance angiography (MRA) for the diagnosis of pulmonary embolism during pregnancy.

Search methods

We searched MEDLINE and Embase until July 2015. We used included studies as seeds in citations searches and in 'find similar' functions and searched reference lists. We approached experts in the field to help us identify non-indexed studies.

Selection criteria

We included consecutive series of pregnant patients suspected of pulmonary embolism who had undergone one of the index tests (computed tomography (CT) pulmonary angiography, lung scintigraphy or MRA) and clinical follow-up or pulmonary angiography as a reference test.

Data collection and analysis

Two review authors performed data extraction and quality assessment. We contacted investigators of potentially eligible studies to obtain missing information. In the primary analysis, we regarded inconclusive index test results as a negative reference test, and treatment for pulmonary embolism after an inconclusive index test as a positive reference test.

Main results

We included 11 studies (four CTPA, five lung scintigraphy, two both) with a total of 695 CTPA and 665 lung scintigraphy results. Lung scintigraphy was applied by different techniques. No MRA studies matched our inclusion criteria.

Overall, risk of bias and concerns regarding applicability were high in all studies as judged in light of the review research question, as was heterogeneity in study methods. We did not undertake meta-analysis. All studies used clinical follow-up as a reference standard, none in

a manner that enabled reliable identification of false positives. Sensitivity and negative predictive value were therefore the only valid test accuracy measures.

The median negative predictive value for CTPA was 100% (range 96% to 100%). Median sensitivity was 83% (range 0% to 100%).

The median negative predictive value for lung scintigraphy was 100% (range 99% to 100%). Median sensitivity was 100% (range 0% to 100%).

The median frequency of inconclusive results was 5.9% (range 0.9% to 36%) for CTPA and 4.0% (range 0% to 23%) for lung scintigraphy. The overall median prevalence of pulmonary embolism was 3.3% (range 0.0% to 8.7%).

Authors' conclusions

Both CTPA and lung scintigraphy seem appropriate for exclusion of pulmonary embolism during pregnancy. However, the quality of the evidence mandates cautious adoption of this conclusion. Important limitations included poor reference standards, necessary assumptions in the analysis regarding inconclusive test results and the inherent inability of included studies to identify false positives. It is unclear which test has the highest accuracy. There is a need for direct comparisons between diagnostic methods, including MR, in prospective randomized diagnostic studies.

PLAIN LANGUAGE SUMMARY

Scanning imaging techniques for the exclusion of pulmonary embolism during pregnancy

Pulmonary embolism is a blood clot that blocks blood flow to a portion of the lungs. Pregnant women are at high risk of pulmonary embolism, and it is a leading cause of death during pregnancy. Women at risk are treated with blood thinning medication. It is important that no cases are missed, and that treatment is prevented in women without the disease. Pulmonary embolism can be diagnosed through different scanning techniques. Little is known about the performance of these tests during pregnancy, which might be different from their performance outside pregnancy. We performed this review to establish the accuracy of the following imaging tests for diagnosing pulmonary embolism during pregnancy: computed tomography pulmonary angiography, lung scintigraphy and magnetic resonance angiography.

We found 11 studies (current until July 2015) that described 695 computed tomography pulmonary angiography results, 665 lung scintigraphy results and no magnetic resonance angiography results. Studies on lung scintigraphy used varying techniques. Overall, these studies were of poor quality; therefore, we could not analyse results together to obtain a single estimate of their accuracy. The identified studies followed-up patients clinically to confirm the absence of pulmonary embolism as revealed on the initial scan, so information could be used to draw conclusions only on the ability of these imaging tests to exclude pulmonary embolism, not on their ability to establish the diagnosis.

Both computed tomography pulmonary angiography and lung scintigraphy appear appropriate for excluding pulmonary embolism in pregnancy. Almost no cases were missed, especially when the imaging test indicated the absence of disease without a doubt. However, this result should be interpreted with care because of the low quality of and variation between identified studies. Around 5% of the scans were unclear, but this percentage was as high as 36% in one study. About 3% of all women included in the studies had pulmonary embolism. We do not know which of the tests is better because tests were not directly compared in the same patients, and because aspects besides test accuracy need to be taken into account. Major limitations of this review include the use of clinical follow-up within studies to confirm the absence of disease, unclear test results and the inability of studies to provide information on the accuracy of these tests in establishing rather than rejecting the diagnosis. High-quality research is needed to investigate the use of computed tomography pulmonary angiography, lung scintigraphy and magnetic pulmonary angiography in the same patient groups.

SUMMARY OF FINDINGS

Summary of findings 1.   Summary of findings table

What is the diagnostic accuracy of imaging tests for the diagnosis of pulmonary embolism (PE) in pregnancy?						
Patients	Pregnant women with clinical suspicion of PE.					
Prior testing and prevalence	Varied. Most often performed were chest X-ray and imaging for deep venous thrombosis. The median prevalence of PE was 3.3% (range 0.0% to 8.7%), as assessed by the applied reference standard, which has limitations.					
Settings	Secondary and tertiary care, both inpatients and outpatients.					
Index test	Computed tomography pulmonary angiography (CTPA), lung scintigraphy and magnetic resonance angiography (MRA). No studies on MRA were included. Inconclusive test results were regarded as negative in the primary analysis.					
Importance	Pregnant women are often suspected of PE because of increased risk and physiological signs that mimic symptoms of PE. Pregnant women are often excluded from diagnostic imaging studies. These imaging tests might perform differently during pregnancy, and radiation and other risks are weighed differently.					
Reference standard	Clinical follow-up of at least 6 weeks. In almost all studies, follow-up was performed to identify PE, not to exclude it. Pulmonary angiography was preferred but was applied by none of the studies.					
Studies	Cross-sectional cohort studies were included. Case-control studies were excluded. All studies were retrospective.					
Test	Number of studies (number of index test results)	Median negative predictive value (range)	Median sensitivity (range)	Median inconclusive test results (range)	Overall risk of bias (QUADAS-2)	Overall applicability (QUADAS-2)
CTPA	6 (695)	100% (96%-100%)	83% (0%-100%)	5.9% (0.9%-36%)	High risk	High concern
Lung scintigraphy	7 (665)	100% (99%-100%)	100% (0%-100%)	4% (0%-23%)	High risk	High concern
CAUTION: The results in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review.						

CTPA: computed tomography pulmonary angiography.
MRA: magnetic resonance angiography.
PE: pulmonary embolism.

BACKGROUND

Target condition being diagnosed

Pulmonary embolism (PE) is a leading cause of maternal mortality (Chang 2003; Kobayashi 2008; Lewis 2007). Approximately 30% of maternal deaths directly related to pregnancy between 2003 and 2005 in the United Kingdom were caused by thromboembolism (Lewis 2007). The incidence approximates 3 per 10,000 pregnancies, including the postpartum period, around 30% of which occur before the time of delivery (Meng 2015). Mortality in untreated patients with PE is reported to be as high as 18% to 35%, although these figures are based on older studies (Calder 2005). Treatment of women with pregnancy-related PE consists of subcutaneous low-molecular-weight heparin (LMWH) (Bates 2012). Drawbacks of LMWH treatment include the need for daily injections, the occurrence of skin reactions and bruising in 2% to 29% of women (Bank 2003; Greer 2005), restricted epidural analgesics during labor and theoretical risks of heparin-induced thrombocytopenia and major bleeding.

Pulmonary embolism together with deep venous thrombosis (DVT) is considered to be one disease entity, and this combination is referred to as venous thromboembolism (VTE). These conditions often co-occur and share the same risk factors and pathophysiology. Pregnant women have a five-fold increased risk of VTE compared with age-matched controls (Pomp 2008). However, absolute numbers of pregnant women with clinical suspicion of PE are still relatively low. Most physicians working in the field of thrombosis infrequently encounter pregnant women with PE, thus experience is greatly lacking. A major challenge is that many symptoms of PE are similar to physiological manifestations of pregnancy, such as swollen legs, dyspnoea and tachycardia. Among pregnant women with suspected PE, a VTE diagnosis is confirmed in 4%, as compared with 12% among non-pregnant patients (Kline 2014). Pulmonary angiography, which traditionally has been considered the gold standard for PE, is an X-ray-based imaging technique that uses contrast agent to visualise thrombi. However, this approach is rarely used anymore because of associated complications, high radiation doses, high costs and declining availability and expertise (Rosenberg 2007; Somarouthu 2010).

Index test(s)

This review focuses on computed tomography pulmonary angiography (CTPA), lung scintigraphy and magnetic resonance angiography (MRA), performed as first-line tests or after other diagnostic testing in pregnant patients suspected of PE. CTPA is an X-ray-based imaging modality that uses intravenous contrast along with computer processing to create tomographic images. CTPA has a clinical validity comparable with that of pulmonary angiography (Quiroz 2005). It is widely available and has the potential of providing an alternative diagnosis. Outside of pregnancy, it is the imaging test most often performed to identify PE. An intraluminal filling defect is diagnostic of PE, and absence of a filling defect rules out the diagnosis. An inconclusive test result is reported for only a small percentage of CTPAs performed outside of pregnancy (0.9% to 4.6% (Mos 2009)). Disadvantages of CTPA include allergic reactions to the contrast agent, risk of contrast-related nephropathy and exposure of both mother and foetus to radiation. The level of foetal exposure seems acceptable provided appropriate methods of dose reduction are applied (Nijkeuter

2006). Concerns have been raised that breast tissue, which is especially susceptible to carcinogenic effects, receives particularly high doses of radiation (Einstein 2007). Iodinated contrast agents cross the placenta and carry the theoretical risk of neonatal thyroid function depression, although this risk is not supported by findings of clinical studies (Bourjeily 2010; Webb 2005).

Lung scintigraphy is another well-established diagnostic modality for PE. Radioisotopes are infused and inhaled to evaluate perfusion and ventilation of the lungs, respectively. This procedure uses gamma cameras to make two-dimensional projection images. Results of ventilation and perfusion scanning are combined to identify a mismatch indicative of PE. Ventilation/perfusion single-photon emission computed tomography, or V/Q SPECT, is based on the same technique but provides three-dimensional imaging. Therefore, V/Q SPECT is in effect a distinct diagnostic modality. It is recommended as a first-line diagnostic procedure by the European Association of Nuclear Medicine guidelines (Bajc 2009). Scan results can be divided into three categories: 'PE present' (high probability), 'PE absent' (very low probability or normal) and 'inconclusive'. Outside pregnancy, the sensitivity of a high-probability lung scintigraphy scan was reported to be 77% and the specificity of a very low probability or normal scan was 98% after non-diagnostic readings were excluded (Sostman 2008). The high percentage (around 39%) of non-diagnostic readings is the main drawback of the V/Q scan and necessitates further imaging (PIOPED 1990). This proportion can be reduced by performing V/Q scan only in patients with a normal chest X-ray. A normal chest X-ray can also justify leaving out the ventilation scan, thereby limiting radiation exposure (Torbicki 2014).

The mother is generally exposed to less radiation with lung scintigraphy than with CTPA (Hurwitz 2006). Radiation exposure of the foetus during lung scintigraphy is equal to or greater than that reported in CTPA, depending on the CT scanner model and imaging protocol used, as well as the trimester and method of estimated exposure (Doshi 2008; Hurwitz 2006a; Winer-Muram 2002). The main risk of in utero exposure to radiation is induction of malignancy. The number of cases of excess malignancy up to age 15 years following irradiation in utero is considered to be 1 in 16,000 per mSv (Streffer 2003). CTPA exposes the foetus to 0.013 mSv of radiation, and a perfusion scan to 0.11 to 0.20 mSv (Nijkeuter 2004). Given the difficulty of calculating radiation and variation in the protocols used, these numbers should be interpreted as equally low. The small amounts justify the use of radiation to prevent a potentially life-threatening PE. Of course, radiation exposure should be kept as low as possible, in line with the As Low As Reasonably Achievable principle (Hendee 1986). At the same time, clinicians must be aware that risks of radiation should be weighed against risks of untreated PE or complications of unnecessary anticoagulant therapy.

Magnetic resonance angiography (MRA) applies magnetic fields and radiofrequency pulses to visualise blood vessels using flow effects or contrast material. This approach offers the advantage that it is radiation-free. However, MRA is not applied in daily practice because of long acquisition times, and few studies have reported on this technique (Kanal 1992; Oudkerk 2002). Another concern is the safety of gadolinium-containing contrast agents during pregnancy (Chen 2008). No deleterious effects have been documented during human pregnancy, but supraclinical doses show deleterious effects on both pregnancy and foetal development in animal studies

(Tremblay 2012). Progress has been made with contrast-free magnetic resonance imaging (Kluge 2004). Although magnetism and radiofrequency pulses carry theoretical risks for the foetus, limited research has produced no evidence of harmful effects on the (unborn) child or on pregnancy (Tremblay 2012).

Clinical pathway

In practice, the diagnostic pathway in the case of suspicion of PE in pregnant women may differ among different physicians and different patients. Depending on the clinical presentation, a chest X-ray or an electrocardiogram can be used to rule out other diseases. As PE and DVT are manifestations of the same disease entity (VTE), physicians sometimes opt for compression ultrasonography (CUS) of the legs in cases of suspected PE, when symptoms of DVT are also present. When the diagnosis of DVT is confirmed, no further imaging of the lungs is necessary, as treatment for patients with PE is the same as treatment for patients with DVT.

For the diagnosis of PE outside of pregnancy, different tests and clinical decision rules can be combined to provide the safest diagnostic strategies. Clinical decision rules are used to determine a patient's pretest probability of PE. In non-pregnant patients, a low to moderate pretest probability combined with a normal D-dimer test is safe to rule out PE without the need for further diagnostic imaging. In pregnancy, however, several items of these rules constitute normal physiological manifestations of pregnancy, for example, swollen legs or heart rate above 100 beats/min. Other items are seldom seen in pregnancy, such as the presence of cancer. Clinical decision rules for PE have not been validated in pregnant women. In addition, during normal pregnancy, D-dimer levels gradually rise. A normal D-dimer level is infrequently encountered in the second and third trimesters of pregnancy (Chabloz 2001; Quiroz 2005); therefore, current practice seldom includes D-dimer levels or clinical decision rules in the diagnostic strategy for ruling out PE during pregnancy. We have not included these tests in this review.

Rationale

It is important to regard pregnant women as a distinct subgroup for which tests might perform differently and additional risks need to be considered. Other research groups have reviewed the available literature on PE diagnostics specifically for pregnant women (Bourjeily 2010a; Brown 2010; Duran-Mendicuti 2011; Nijkeuter 2006; Rodger 2010; Rosenberg 2007; Tan 2011). These narrative reviews are useful as guides in the clinical dilemma of suspected PE in pregnancy. However, they lack the rigorous and explicit methods of a Cochrane review with transparent guidelines, minimisation of bias and periodic updates.

An objectively confirmed diagnosis of PE is imperative in pregnancy. Anticoagulant treatment is potentially harmful and has implications for delivery, and a PE diagnosis determines future decisions on thromboprophylaxis and oral contraceptive use. On the other hand, untreated PE is a potentially fatal disease. Physiological changes during pregnancy may mimic the symptoms of PE; therefore, PE can be suspected more often in pregnant women. D-dimer testing and clinical decision rules available for non-pregnant patients are not useful for pregnant women.

Diagnostic imaging tests for PE have been extensively studied in non-pregnant patients (Bounameaux 2010), but pregnant women usually are excluded from these diagnostic and management studies. A strong evidence base for the management of pregnancy-related PE is therefore missing (Middeldorp 2011). Extrapolating results from studies and guidelines on non-pregnant patients is not advisable because pregnant women form a clinically distinct subgroup. The average age of pregnant women is lower than the age of patients in the studies from which diagnostic accuracy data are derived. Therefore, pregnant women are less likely to have cardiorespiratory comorbidity, which can result in abnormal V/Q scans (Matthews 2006). In pregnancy, a raised diaphragm might result in abnormal perfusion scan results.

Another issue complicating the diagnosis of PE specifically in pregnancy is that imaging tests expose the foetus to ionising radiation. Although the level of foetal radiation exposure is far below the threshold for inducing carcinogenic and teratogenic effects, physicians may be hesitant to perform imaging tests. Given the high index of suspicion and the dangers of misdiagnosis encouraging diagnostic efforts, along with radiation exposure and other test complications discouraging testing, optimal diagnostic strategies are even more crucial in pregnant women than in non-pregnant patients.

O B J E C T I V E S

To determine the diagnostic accuracy of CTPA, lung scintigraphy and MRA for the diagnosis of PE during pregnancy. and to calculate the diagnostic accuracy of these tests in relation to the reference standard of pulmonary angiography or clinical follow-up for at least six weeks for occurrence of symptomatic PE.

Secondary objectives

To determine the number of inconclusive test results for each diagnostic test or diagnostic strategy, defined as neither showing nor excluding PE.

We aimed to investigate the effects of following clinical factors on the negative predictive value of index tests.

1. Prior testing (e.g. using clinical decision rule results and D-dimer testing).

2. Gestational age.

3. Type of reference standard.

4. Technological advances of CTPA (multi-slice vs single-slice).

M E T H O D S

Criteria for considering studies for this review

Types of studies

We considered studies with a prospective or retrospective consecutive patient series design eligible for inclusion. The minimum sample size was set at 25 per index test, in light of the a priori chance of PE of 4%; we deemed smaller studies uninformative. Studies could still be included if patients received one of multiple index tests and the choice between tests was not random and was not based on a clear protocol, as long as recruitment was consecutive. We considered inclusion consecutive

if study authors described it using this or similar terms. The quality item on consecutive inclusion could still be scored 'no' on the basis of criteria provided in Table 1 (see below). In studies in which the index test could be performed multiple times for one patient, study authors reported the index test result on a patient or scan basis. In both cases, data on index test results in relation to reference results had to be extractable from the article or provided on request by study authors. We excluded case-control studies.

Participants

Participants were pregnant women with clinical suspicion of PE. Conditions for suspicion of PE and confirmation of pregnancy were left to the discretion of the authors of included studies. We did not include data on postpartum or asymptomatic patients.

If not all women in the study met the inclusion criteria of the present review, data had to be extractable for relevant participants. If relevant data were not extractable and could not be supplied by the study author, we excluded the study. Study populations from secondary and tertiary care settings, comprising inpatients and outpatients, were eligible for inclusion. We excluded primary care patients because test accuracy can vary with care setting.

Index tests

Studies investigating CTPA, lung scintigraphy, either planar or V/Q SPECT and either with or without ventilation scanning, and MRA as an index test were eligible for inclusion. If the patient had undergone the same index test more than once, we included only data on the first index test performed. If the first test was followed by a different index test, we included this test result. Data on original assessments of the index test result in the clinical setting were preferred over re-assessments for study purposes.

Target conditions

The target condition of this review is symptomatic PE. For studies that provided multiple definitions of non-diagnostic scans according to the level of the pulmonary vasculature, we considered scans as non-diagnostic when they were of insufficient quality to allow visualisation of PE at segmental or more proximal levels.

Reference standards

We accepted as reference standards both pulmonary angiography and clinical follow-up for the occurrence of symptomatic PE for at least six weeks. We anticipated that few or no studies would use pulmonary angiography as the reference standard. We considered a PE confirmed within three months of the index test as a positive reference test.

The diagnosis of PE was confirmed on pulmonary angiography in the case of a new intraluminal filling defect, with cut-off of contrast material in a vessel with diameter greater than 2.5 mm.

If clinical follow-up was applied as the reference standard, objective diagnostic testing had to be performed in the case of symptoms of PE. Objective criteria for the diagnosis of PE included a (new) intraluminal filling defect on CTPA; cut-off of contrast material in a vessel greater than 2.5 mm in diameter; a new perfusion defect involving at least 75% of a segment, with corresponding normal ventilation (i.e. a high-probability lung scan); a new non-diagnostic lung scan accompanied by documentation of DVT on ultrasonography or venography; and confirmation of a new PE

at autopsy. For clinical follow-up, PE was also diagnosed when symptoms of PE were combined with a diagnosis of DVT. The objective criterion for the diagnosis of DVT was a (new) non-compressible venous segment or a substantial increase (4 mm or more) in the diameter of the thrombus during full compression in a previously abnormal segment on ultrasonography, or a new intraluminal filling defect on venography. In case of an inconclusive index test, we considered the reference test positive if the patient was treated on clinical grounds without objective evidence of VTE on follow-up. This provides a conservative estimate of both negative predictive value and sensitivity in the primary analysis.

When both pulmonary angiography and clinical follow-up were performed, we considered pulmonary angiography the reference standard.

Search methods for identification of studies

We applied no language restrictions.

Electronic searches

We searched MEDLINE and Embase (OvidSP) systematically for potentially eligible studies until July 2015. We specified the search strategies used in Appendix 1 and Appendix 2. We used three sets of terms reflecting the research question of the review: the tests (index tests), the target condition (PE) and the patients (pregnant women). We executed searches with venous thrombosis or thromboembolism as the target condition to increase sensitivity. We did not use a methodological search filter or a search term such as 'diagnostic' because doing this has been shown to decrease the sensitivity of the search without significantly benefiting the specificity (Leeflang 2006; Ritchie 2007).

Searching other resources

We used studies identified by the electronic search as eligible for inclusion in the review as 'seeds' for the 'related article' feature in PubMed and the 'find similar' feature in Embase to identify other potentially relevant articles. We performed a citations search for these studies in Web of Science to identify articles that cited them. We handsearched the reference lists of included studies and of previous reviews on this topic for potentially relevant studies. Finally, we contacted experts in the field (MV Huisman, M Righini, WS Chan) to enquire about unpublished studies and studies not included in Embase or MEDLINE.

Data collection and analysis

Selection of studies

Review authors retrieved articles identified by the search and reviewed all abstracts of potentially relevant studies, judging them first by title with a very low threshold. We then conducted a full-text review of publications considered potentially relevant on the basis of the abstract review to determine whether they met the inclusion criteria defined above. Two review authors independently completed this process. We resolved disagreements by discussion or, when necessary, through involvement of a third review author, whose judgement was then decisive.

Data extraction and management

Two review authors used standardised, piloted data extraction forms to independently extract data from all selected studies. We resolved disagreements by discussion or, when necessary,

through involvement of a third review author, whose judgement was then decisive. Review authors collected the following data from each study for each test: general information (title, study authors, year of publication, status of publication, period of recruitment, study design); descriptives of the relevant patient population (age, sex, ethnicity, country of recruitment, inpatients or outpatients, healthcare setting, PE symptoms); information on the test under study; reference test used (clinical follow-up or pulmonary angiography, or both); items for study quality assessment (see below); data on intraobserver and interobserver variability; numbers of inconclusive test results in diseased and non-diseased groups; and numbers of true and false positives and true and false negatives for relevant women. We contacted the authors of the original studies in cases of missing data.

Assessment of methodological quality

Two review authors independently assessed the methodological quality of selected studies using the QUADAS-2 instrument (Table 1) - an evidence-based tool for assessment of the methodological quality of diagnostic studies included in systematic reviews (Whiting 2011). Review authors scored each item 'yes', 'no' or 'unclear' according to the criteria outlined in Table 1 and provided additional judgements when quality issues fell outside the scope of the criteria. We resolved disagreements by discussion or, when necessary, with involvement of a third review author, whose judgement was then decisive.

Statistical analysis and data synthesis

From each study, we extracted available data on true and false negatives and positives for relevant participants to calculate specificity, sensitivity and post-test probabilities. We planned to pool the data in a meta-analysis for each diagnostic test if a sufficient number of studies of sufficient quality and methodological uniformity were available. For PE prevalence, we divided true positives plus false negatives by total sample size.

We intended to use a bivariate random-effects model in calculating summary estimates of sensitivity and specificity (Reitsma 2005) and to translate summary estimates into likelihood ratios. To illustrate the findings, we intended to present post-test probabilities for a range of plausible values of pretest probabilities of PE.

Besides the possibility of being positive or negative, a test might be inconclusive. In practice, patients with an inconclusive test result often undergo further testing. The conventional approach to calculating sensitivity and specificity does not provide a way of incorporating this third outcome category. Ignoring the inconclusive category would lead to overestimations of sensitivity and specificity. Calculating the sensitivity by dividing true positives by the total number of participants with the disease suggests that 100% minus that percentage is the proportion of diseased participants missed by the test. However, this is not the case, because this proportion also comprises participants who did not have a conclusive test result. Even though they did not have a conclusive test result, these participants are not necessarily falsely classified. Calculating specificity for a test with an inconclusive result category is similarly thwarted. For index tests with inconclusive results, we explicitly stated the proportions of participants with an inconclusive test result in the diseased group and in the healthy group, respectively. In our primary analyses, we calculated test accuracy measures while considering intermediate results as 'negatives', with the aim of providing conservative

estimates rather than overestimations. In sensitivity analyses, we re-analysed the data while treating inconclusive results as 'positives'.

Clinicians may be especially interested in the proportion of patients with PE that will be missed after a negative index test, or after an inconclusive index test. Therefore, we also estimated summary post-test probabilities of PE for the three index test result categories if meta-analysis provided summary estimates for sensitivity and specificity. This could be done by extending the bivariate model for multiple index test categories, provided sufficient data were available, and provided the models converged (Bipat 2007; Leeflang 2012).

Investigations of heterogeneity

To explore anticipated sources of heterogeneity, we intended to add predefined covariates to the meta-regression. We would undertake this only for index tests or strategies that were investigated in more than three studies reporting on the particular covariate (Higgins 2009; Lambert 2002; Morton 2004; Thompson 2002).

Clinical factors included the following.

1. Prior testing and clinical decision rule results. Patients with suspected PE are often stratified into high- and low-risk groups on the basis of clinical decision rules or empirical clinical likelihood. The prevalence of PE differs between these groups. Results of prior diagnostic tests also influence the likelihood that the target condition is present. Therefore, we would record and introduce into the analysis of heterogeneity any clinical decision rule or other prior diagnostic test that was not the index test of the particular study.
2. Stage of pregnancy. For various reasons, pregnant women should be considered a distinct subgroup with regard to the diagnosis of PE. The importance of these reasons can vary during the course of pregnancy, for example, the progressive rise in D-dimer levels or signs of pregnancy that mimic symptoms of PE.
3. Reference standard. It is desirable to allow only one reference standard for all studies. Pulmonary angiography used to be the gold standard, but in more recent literature, it has been replaced by clinical follow-up. Therefore, we decided to include both as acceptable reference standards. This might lead to differential verification and may cause heterogeneity.
4. Technological advances of index tests: multi-slice versus single-slice CTPA.

We intended to also explore heterogeneity more informally by visually inspecting forest plots stratified by potential factors of heterogeneity.

Sensitivity analyses

We planned to perform sensitivity analysis to explore effects on overall results of historical versus current inclusion of participants. Diagnostic accuracy studies are by definition cross-sectional because two tests are compared on the basis of their ability to provide a diagnosis at the same moment in the course of a disease. However, inclusion of patients can be historical, that is, retrospective. This design is more prone to selection bias and is therefore an anticipated source of heterogeneity.

We also investigated the effect of considering inconclusive index test results as 'positive' and compared these results with when they were considered to be 'negative'.

Assessment of reporting bias

We did not explore publication bias because adequate methods of detecting publication bias have not yet been developed for use in diagnostic accuracy studies (Deeks 2005).
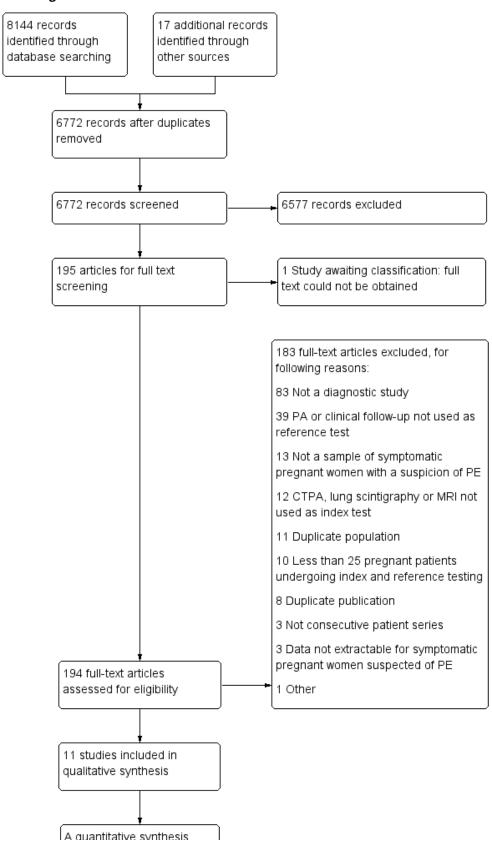
R E S U L T S

Results of the search

See Figure 1.

Figure 1. Study flow diagram.

Figure 1. Study flow diagram.

Figure 1.   (Continued)

> A quantitative synthesis
> (meta-analysis) could not
> be performed

Our search yielded 5841 hits in Embase and 2303 hits in MEDLINE. We identified another 17 articles through handsearching of references. Other elements of the search strategy revealed no additional potentially eligible studies. After duplicates were removed, 6772 publications remained for title and abstract screening. We deemed a total of 195 as potentially eligible and marked them for full-text review. Of these, we excluded 183 for reasons listed in Figure 1. We could not obtain the full text of one study (Vin 2009). We included the remaining 11 studies in the present review (Bajc 2015; Bourjeily 2012; Browne 2014; Chan 2002; Cutts 2014; Hullah 2011; Litmanovich 2009; Nijkeuter 2013; Ridge 2009; Scarsbrook 2007; Shahir 2010). The original authors of eight studies provided additional data or clarification on their study methods, as did study authors for three studies that we eventually excluded. We identified 11 studies that presented data on a study population that was also reported in another publication.

Characteristics of studies

We provided characteristics and quality assessments of individual included studies in the Characteristics of included studies table. We listed selected excluded studies that readers might consider in light of the research question, or that appeared on first inspection to be eligible, in the Characteristics of excluded studies table, along with reasons for exclusion. We listed in Figure 1 reasons for exclusion of all 183 studies that we excluded after full-text review. We identified only one potentially eligible study on MRA (Herédia 2012) but excluded this study because of its small sample size (n < 25).

Of the included studies, five reported on lung scintigraphy (Bajc 2015; Chan 2002; Cutts 2014; Hullah 2011; Scarsbrook 2007), four on CTPA (Bourjeily 2012; Browne 2014; Litmanovich 2009; Nijkeuter 2013) and two on both lung scintigraphy and CTPA (Ridge 2009; Shahir 2010). CTPA was also reported in Hullah 2011 and Scarsbrook 2007 but their sample size was too small. The total numbers of index test results that we could include in the current review were 695 for CTPA and 665 for lung scintigraphy.

All studies were retrospective consecutive patient series that used clinical follow-up as a reference standard. Not all were designed with the aim of determining the diagnostic accuracy of index tests. Other aims included evaluating image quality (Browne 2014), testing diagnostic validity of clinical prediction models (Cutts 2014) and assessing radiation dose (Litmanovich 2009).

Of the CTPA studies, one used single-detector or multi-detector scanners (Nijkeuter 2013), and the others used only multi-detector scanners (Bourjeily 2012; Browne 2014; Litmanovich 2009; Ridge 2009; Shahir 2010). Of the lung scintigraphy studies, one used SPECT and performed ventilation scanning in a proportion of patients (Bajc 2015). The other studies used planar scintigraphy. Two studies performed V/Q scanning for all patients (Chan 2002; Hullah 2011), and two performed only perfusion scanning (Scarsbrook 2007; Shahir 2010). All patients in these studies had

undergone a prior chest X-ray. One study did not specifically state whether ventilation scanning was performed (Cutts 2014). One lung scintigraphy study performed ventilation scanning in a minority of patients (Ridge 2009).

For studies reporting the duration of follow-up, median duration was six months (Browne 2014; Chan 2002; Litmanovich 2009; Nijkeuter 2013; Ridge 2009; Scarsbrook 2007; Shahir 2010). Two studies provided follow-up throughout pregnancy and the postpartum period (Bajc 2015; Cutts 2014); one over the postpartum period or three months, whichever was longer (Bourjeily 2012) and one until delivery (Hullah 2011).

Recruitment starting year ranged from 1990 to 2009. The recruitment end year ranged from 2000 to 2013. Countries of inclusion were USA (n = 3), Australia (n = 2), Ireland (n = 2), United Kingdom (n = 2), Canada (n = 1), the Netherlands (n = 1) and Sweden (n = 1). Four studies recruited only at tertiary care centres (Bajc 2015; Bourjeily 2012; Cutts 2014; Hullah 2011), one at a secondary care centre (Ridge 2009) and three at both types of centres (Chan 2002; Nijkeuter 2013; Shahir 2010). For the remaining three studies, this information was unclear. Two studies recruited patients from the emergency department (Litmanovich 2009; Ridge 2009), and three recruited both inpatients and outpatients (Bajc 2015; Chan 2002; Nijkeuter 2013). For the remaining six studies, this information was unclear.

The median of reported mean or median age was 31 years (range 29 to 32.6 years). Two studies did not report gestational age (Bourjeily 2012; Ridge 2009). In one study, mean gestational age was 19 weeks (Litmanovich 2009). In the other studies, a median of 9% of patients presented in the first trimester, a median of 41% in the second trimester and a median of 49% in the third trimester.

Four studies described presenting symptoms (Bajc 2015; Bourjeily 2012; Chan 2002; Shahir 2010). In all four studies, the most common symptom was dyspnoea, and the second most common was chest pain.

Nine studies described some extent of prior or additional testing (Bajc 2015; Browne 2014; Chan 2002; Cutts 2014; Hullah 2011; Litmanovich 2009; Ridge 2009; Scarsbrook 2007; Shahir 2010). A total of 230 patients underwent imaging for DVT. One case of DVT was found. Other often performed tests included D-dimer and chest X-ray. Four studies reported that a small proportion of up to six patients underwent the index test multiple times or received a combination of lung scintigraphy and CTPA (Browne 2014; Hullah 2011; Ridge 2009; Shahir 2010).

Methodological quality of included studies

We summarised risk of bias and applicability concerns in Figure 2 and Figure 3. None of the studies had the main aim of estimating specificity of the index test, and only some were designed to estimate the sensitivity or negative predictive value (Bajc 2015;

Bourjeily 2012; Nijkeuter 2013; Ridge 2009; Scarsbrook 2007). We assessed quality in terms of risk of bias and concerns regarding applicability in light of the research question of the current review, not the research questions of individual studies. From this perspective, the overall quality of the included studies was low. We deemed that no studies were low risk, and that none produced a low level of concern on more than three items of the QUADAS-2 tool. The item that had lowest risk of bias overall was the index test. We scored a relatively high proportion of items as unclear, in part because of the divergence between the research question of the current review and the research questions of individual studies.

Figure 2. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.
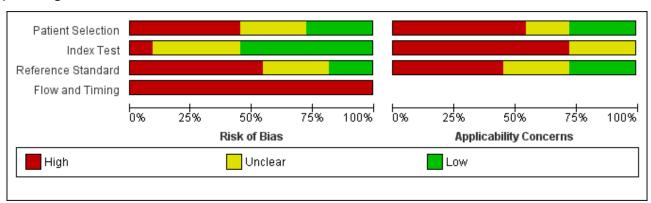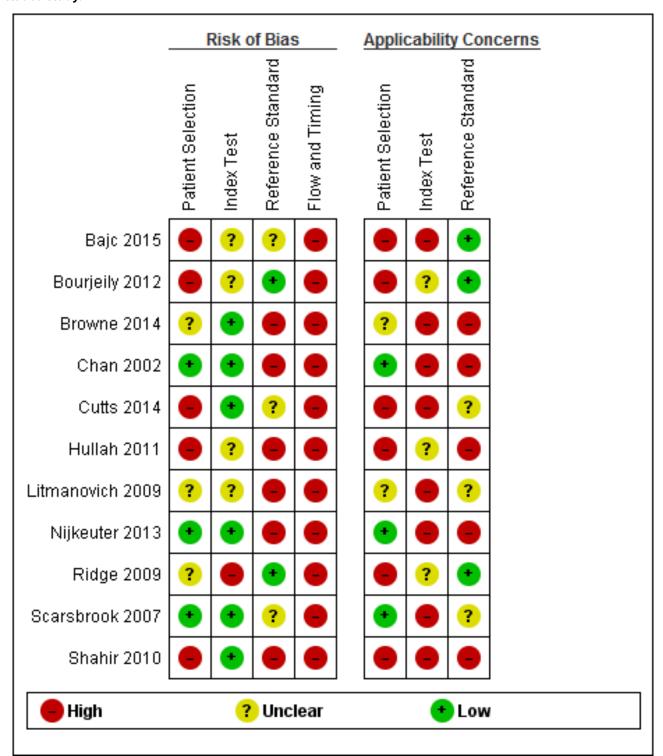
Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

Patient selection

We judged patient selection to carry low risk of bias in three studies (Chan 2002; Nijkeuter 2013; Scarsbrook 2007), unclear risk in three studies (Browne 2014; Litmanovich 2009; Ridge 2009) and high risk in five studies (Bajc 2015; Bourjeily 2012; Cutts 2014; Hullah 2011; Shahir 2010). Concern about applicability was low in three studies (Chan 2002; Nijkeuter 2013; Scarsbrook 2007), unclear in two studies (Browne 2014; Litmanovich 2009) and high in six studies (Bajc 2015; Bourjeily 2012; Cutts 2014; Hullah 2011; Ridge 2009; Shahir 2010). Causes for high concern and high risk of bias included non-systematic selection between index tests (Bajc 2015; Hullah 2011; Ridge 2009; Shahir 2010), anticoagulant treatment that was

started before testing was begun in over 5% of patients (Hullah 2011; Shahir 2010), a suboptimal patient selection procedure for a consecutive series of pregnant women (Cutts 2014; Hullah 2011) and high prevalence of comorbidity (Bourjeily 2012).

Index test

We judged the risk that conduct or interpretation of the index test had introduced bias as low in six studies (Browne 2014; Chan 2002; Cutts 2014; Nijkeuter 2013; Scarsbrook 2007; Shahir 2010), high in one (Ridge 2009) and unclear in the remaining four (Bajc 2015; Bourjeily 2012; Hullah 2011; Litmanovich 2009). Among low-risk studies, interpretation of the index test was blinded, but in the other studies this was unclear. In none of the studies was concern regarding applicability low. Concern was high in eight studies (Bajc 2015; Browne 2014; Chan 2002; Cutts 2014; Litmanovich 2009; Nijkeuter 2013; Scarsbrook 2007; Shahir 2010)and unclear in three (Bourjeily 2012; Hullah 2011; Ridge 2009). Reasons for high concern included use of multiple index test procedures (Nijkeuter 2013; Ridge 2009; Scarsbrook 2007; Shahir 2010), re-assessment of index test results by researchers for study purposes (Browne 2014; Chan 2002; Cutts 2014; Litmanovich 2009) and use of SPECT scintigraphy (Bajc 2015).

Reference standard

We judged the risk that conduct or interpretation of the reference test had introduced bias as low in two studies (Bourjeily 2012; Ridge 2009), high in six studies (Browne 2014; Chan 2002; Hullah 2011; Litmanovich 2009; Nijkeuter 2013; Shahir 2010) and unclear in the remaining three studies (Bajc 2015; Cutts 2014; Scarsbrook 2007). Concern regarding applicability was high in five studies (Browne 2014; Chan 2002; Hullah 2011; Nijkeuter 2013; Shahir 2010), low in three studies (Bajc 2015; Bourjeily 2012; Ridge 2009) and unclear in three studies (Cutts 2014; Litmanovich 2009; Scarsbrook 2007). Methods of clinical follow-up varied from a query regarding the hospital radiology system used for any imaging for PE after the index test (Browne 2014) to extensive medical record and radiology system review in all hospitals in the region plus telephone interviews and contact with the general practitioner (Chan 2002). Seven studies used a reference standard that could not be used to identify false positives (Bajc 2015; Bourjeily 2012; Browne 2014; Chan 2002; Litmanovich 2009; Nijkeuter 2013; Shahir 2010), mainly because follow-up was specifically targeted to find evidence of VTE. None of these studies aimed to investigate false positives. Three studies used a reference standard that had the potential to identify false positives (Cutts 2014; Hullah 2011; Ridge 2009), and in one study, this was unclear (Scarsbrook 2007). No studies stated that interpretation of the reference test was blinded.

Flow and timing

We deemed all studies to have high risk of bias in the domain of flow and timing. All studies used clinical follow-up as the reference standard. During the follow-up period there is the possibility that the PE diminishes through the natural course of the disease in case of a PE that was not identified by the index test, or that it occurs in a patient that did not have a PE at the time of the index test. Therefore, clinical follow-up can introduce bias in the domain of flow and timing. In two studies, investigators stated that only cases with a negative or inconclusive index test underwent the reference standard (Chan 2002; Nijkeuter 2013), whereas others did

not clearly describe this. Among studies in which patients could undergo different imaging tests for PE, the criteria used to choose between tests were unclear or non-standardised in four (Bajc 2015; Hullah 2011; Ridge 2009; Shahir 2010) and were standardised in one (Scarsbrook 2007). In total, 20 patients were described as lost to follow-up (Bourjeily 2012; Shahir 2010).

Findings

We deemed meta-analysis of results of individual studies inappropriate because of heterogeneity and overall low methodological quality. We identified numerous sources of heterogeneity. Methods used for patient selection and clinical follow-up varied considerably in thoroughness. Where two index tests were available, determining which one the patient should receive was not performed systematically, (i.e. randomly or on the basis of a clear protocol), except in one study (Scarsbrook 2007). Lung scintigraphy studies used different imaging modalities (perfusion scintigraphy, ventilation/perfusion scintigraphy or V/Q SPECT). If a proportion of patients underwent multiple scans, some studies reported on scans as the unit of analysis, whereas others reported on patients, (i.e. a patient was considered false negative, or a single scan was considered false negative). Studies used either the original clinical assessment of the index test or a re-assessment for study purposes. Prior and additional testing performed was unclear in almost all studies. Comorbidity varied between studies, as did blinding to index test results. Almost no study reported cut-offs or definitions for index test results. Follow-up often was not designed to identify false positives and was not provided for all patients in some studies. We could not perform a sensitivity analysis of historical versus current inclusion because all studies were retrospective in design.

None of the studies was designed with the aim of establishing the accuracy of the index test, in terms of sensitivity and specificity, which was the aim of the current review. Clinical follow-up is inherently poor in detecting false positives. If the index test indicates PE, anticoagulant treatment is started and symptoms decline. In the case of a false positive, with no or self-limiting pathology instead of a PE, symptoms also subside. Furthermore, some studies did not follow patients with a positive index test. Most aimed only to identify the occurrence of VTE during follow-up, usually by searching for medical notes or imaging data from re-admissions after the index test. Clearly, false positives are not detected in this way. Therefore, the only reliable measures of test accuracy that can be established with the available data are sensitivity and negative predictive value (NPV). In the primary analysis, inconclusive index test results were regarded as negative.

CTPA

The median NPV of CTPA was 100%. The lowest reported NPV was 96%, and the highest was 100% (Figure 4). Median sensitivity in the CTPA studies was 83%. For one study, we could not estimate sensitivity because there were no cases of PE (Litmanovich 2009). Therefore the median sensitivity does not derive from exactly the same set of studies as the median NPV. Sensitivity ranged from 0% to 100%, as displayed in Figure 5. The study with a sensitivity of 0% had no positive results on 28 CTPA scans, but one of the patients with an inconclusive scan was treated for PE on clinical grounds, and thus was regarded as false negative (Ridge 2009). No studies identified any false-positive cases.

Figure 4. Primary analysis. Negative predictive values (%) with 95% confidence intervals for CTPA with inconclusives regarded as negative.
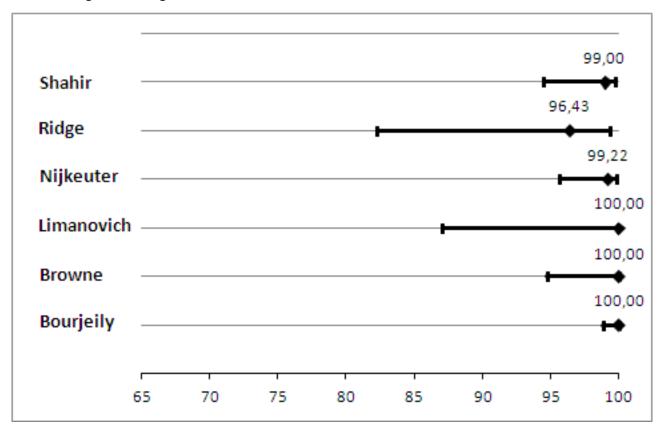

Figure 5. Primary analysis. Forest plot of CTPA with inconclusives regarded as negative.

Study	TP	FP	FN	TN	Inconclusive with PE treatment	Inconclusive without PE treatment	Sensitivity (95% CI)	Sensitivity (95% CI)
Bourjeily 2012	8	0	0	325	0	3	1.00 [0.63, 1.00]	
Browne 2014	1	0	0	69	0	2	1.00 [0.03, 1.00]	
Litmanovich 2009	0	0	0	26	0	3	Not estimable	
Nijkeuter 2013	5	0	1	128	1	7	0.83 [0.36, 1.00]	
Ridge 2009	0	0	1	27	1	9	0.00 [0.00, 0.97]	
Shahir 2010	4	0	1	99	0	6	0.80 [0.28, 0.99]	

The median percentage of inconclusive results was 5.9%, and the percentage reported ranged from 0.9% to 36%. The absolute numbers are displayed in Figure 5. The total number of inconclusive scans was 32. Two women in total had an inconclusive CTPA and were considered to have a positive reference standard. For one woman, lung scintigraphy showed PE (Ridge 2009), whereas the other was treated on clinical grounds (Nijkeuter 2013).

We performed a sensitivity analysis in which we regarded inconclusive index test results as positive. In this analysis, the median NPV was 100%, and NPV ranged from 99% to 100% (Figure 6). Median sensitivity in this analysis was 100%, and sensitivity ranged from 80% to 100% (Figure 7).

Figure 6. Sensitivity analysis. Negative predictive values (%) with 95% confidence intervals for CTPA with inconclusives regarded as positive.
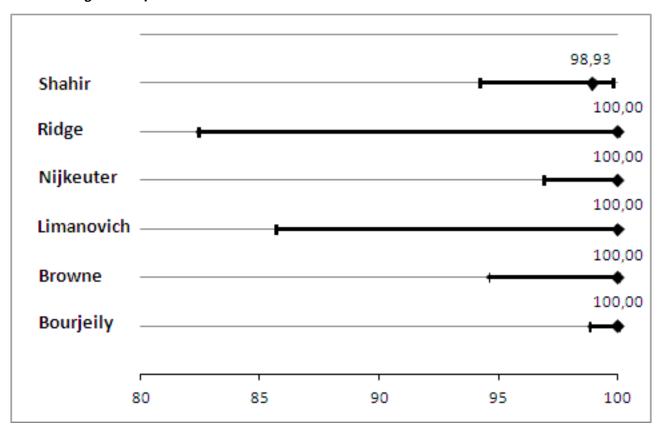

Figure 7. Sensitivity analysis. Forest plot of CTPA with inconclusives regarded as positive.

Study	TP	FP	FN	TN	Inconclusive with PE treatment	Inconclusive without PE treatment	Sensitivity (95% CI)	Sensitivity (95% CI)
Bourjeily 2012	8	3	0	322	0	3	1.00 [0.63, 1.00]	
Browne 2014	1	2	0	67	0	2	1.00 [0.03, 1.00]	
Litmanovich 2009	0	3	0	23	0	3	Not estimable	
Nijkeuter 2013	6	7	0	121	1	7	1.00 [0.54, 1.00]	
Ridge 2009	1	9	0	18	1	9	1.00 [0.03, 1.00]	
Shahir 2010	4	6	1	93	0	6	0.80 [0.28, 0.99]	

Lung scintigraphy

For lung scintigraphy, the median NPV was 100%. The lowest reported NPV was 99%, and the highest was 100% (Figure 8). Median sensitivity was 100%, and sensitivity ranged from 0% to 100%, as displayed in Figure 9. The study with a sensitivity of 0% had no positive results on 97 lung scintigraphy scans, but one of the three patients with an inconclusive scan was found to have PE on a subsequent CTPA, and thus was regarded as false negative (Shahir 2010). No studies identified any false-positive cases.

Figure 8. Primary analysis. Negative predictive values (%) with 95% confidence intervals for lung scintigraphy with inconclusives regarded as negative.
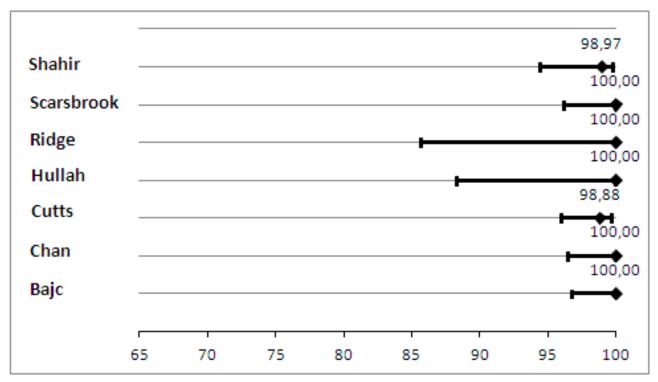

Figure 9. Primary analysis. Forest plot of lung scintigraphy with inconclusives regarded as negative.

The median percentage of inconclusive results was 4.0%, and the percentage reported by study authors ranged from 0% to 23%. The absolute numbers are displayed in Figure 9. The total number of inconclusive scans was 42. Three women in total had an inconclusive lung scintigraphy and were considered to have a positive reference standard. One was an incomplete study because the patient was unable to lie on the scanning table (Shahir 2010). Subsequent CTPA showed PE. The other two patients were treated for PE on clinical grounds (Cutts 2014).

We performed a sensitivity analysis in which we regarded inconclusive index test results as positive. In this analysis, the NPV was 100% in all studies, as shown in Figure 10, reflecting the absence of PE revealed by the reference test after a negative lung scintigraphy. The sensitivity was not estimable in one study because patients receiving PE treatment were excluded and no other PEs were identified by the reference test (Chan 2002). In all other studies, the sensitivity was 100% in this analysis (Figure 11).

Figure 10.   Sensitivity analysis. Negative predictive values (%) with 95% confidence intervals for lung scintigraphy with inconclusives regarded as positive.
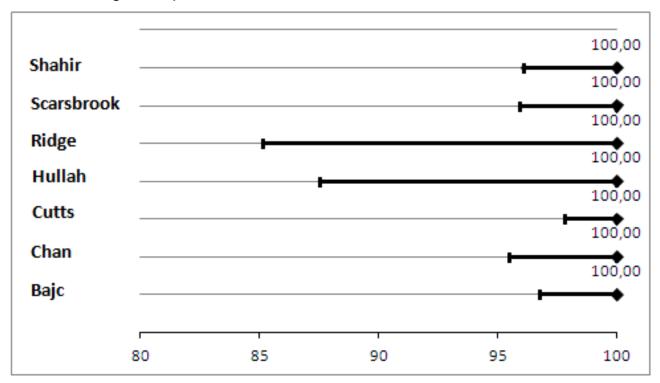
Figure 11.   Sensitivity analysis. Forest plot of lung scintigraphy with inconclusives regarded as positive.
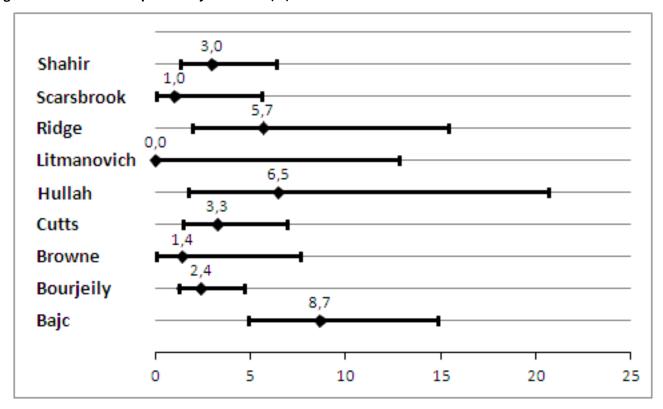
PE prevalence

When we considered all cases with a positive index or reference test as having PE, including treatment after an inconclusive test, we found that the median prevalence in all studies combined was 3.3%, and prevalence ranged from 0.0% to 8.7%, as displayed in Figure 12. We excluded two studies from this analysis because they excluded women receiving therapeutic doses of LMWH (Chan 2002; Nijkeuter 2013).

Figure 12. Prevelance of pulmonary embolism (%) with 95% confidence interval.

DISCUSSION

Summary of main results

The aim of this review was to determine the diagnostic accuracy of imaging tests for the diagnosis of pulmonary embolism (PE) in pregnancy. We included a total of 11 studies reporting on computed tomography pulmonary angiography (CTPA), lung scintigraphy or both. Six studies investigated CTPA (Bourjeily 2012; Browne 2014; Litmanovich 2009; Nijkeuter 2013; Ridge 2009; Shahir 2010), one single-photon emission computed tomography (SPECT) with scintigraphy (Bajc 2015), two planar ventilation/perfusion scintigraphy (Chan 2002; Hullah 2011), two perfusion scintigraphy (Scarsbrook 2007; Shahir 2010) and one perfusion with ventilation scintigraphy in a small proportion of patients (Ridge 2009); another study was unclear on the lung scintigraphy technique used (Cutts 2014). We identified no studies on magnetic resonance angiography (MRA) that met the inclusion criteria of this review. Summary of findings 1 displays a summary of our findings.

All studies were retrospective and used clinical follow-up as the reference standard. Almost all applied clinical follow-up to identify PE, not to exclude it, thus precluding conclusions on test specificity and positive predictive value. No studies identified any false-positive cases. No studies had the aim of determining test specificity, and only some studies were designed to determine sensitivity. We assessed study quality in light of the research question presented in the current review, not in light of research questions posed by the individual studies. From this perspective overall study quality was low and concern regarding applicability was high. We considered meta-analysis inappropriate for both CTPA and lung scintigraphy studies because of heterogeneity and overall low study quality. Sources of

heterogeneity included scintigraphy techniques, thoroughness of clinical follow-up, handling of multiple scans for one patient, use of original assessments or re-assessments for research into index test results, robustness of patient selection, prior testing and blinding.

In six studies reporting a total of 695 CTPA index test results, the median negative predictive value (NPV) was 100%, and NPV ranged from 96% to 100%. CTPA had a median sensitivity of 83%, ranging from 0% to 100%. We found a median of 5.9% inconclusive CTPA test results, ranging from 0.9% to 36%.

In seven lung scintigraphy studies reporting a total of 665 index test results, the median NPV was also 100%, and NPV ranged from 99% to 100%. Median sensitivity was 100%, and sensitivity ranged from 0% to 100%. The median percentage of inconclusive results was 4.0%, with a range of 0% to 23%.

Overall, median PE prevalence was 3.3%, and PE prevalence ranged from 0.0% to 8.7%.

Strengths and weaknesses of the review

Given that the search used broad inclusion criteria without language restrictions and search filters and yielded many screened publications, the chance of missed relevant publications up to the date of the search is small. We attempted to contact all authors of included studies to ensure maximum completeness of study data. By excluding primary care patients and using original assessments of index test results, when available, we optimised the applicability of review findings.

We made methodological choices with the aim of providing conservative estimates rather than overestimations of test

accuracy. There is no proper way to deal with inconclusive test results when determining diagnostic accuracy measures. Excluding them from the analysis leads to overestimation. Artificially regarding them as negative or positive increases specificity or sensitivity, respectively. We regarded inconclusive scans as negative index test results in the primary analysis. Large differences between forest plots in the primary analysis and the sensitivity analysis underline the inherent problem of conventional test accuracy measures when an inconclusive category is involved. Another dilemma is that patients with an inconclusive scan who are treated with anticoagulants on clinical grounds will likely not have a PE identified or excluded during follow-up. This precludes false positives and false negatives in these cases. We regarded treatment after an inconclusive scan as a positive reference standard that provides conservative estimates of NPV and sensitivity.

We judged the overall risk of bias in both CTPA and lung scintigraphy studies as high for the research question of this review. This reflected in part divergence from research questions posed by individual studies, which were not all specifically designed to establish test accuracy. For the clinical research question at hand, one should consider the quality of evidence low and should interpret results cautiously. Major potential causes of bias included non-systematic selection between index tests in four studies (Bajc 2015; Hullah 2011; Ridge 2009; Shahir 2010), unclear blinding to index test results in all studies and use of clinical follow-up.

Clinical follow-up has inherent limitations as a reference standard. First, follow-up is not strictly cross-sectional, so PE may occur or resolve between the time of the index test and the time of the reference test. We chose an upper cut-off of three months between the index test and occurrence of a PE during follow-up. Second, in studies with unscrupulous follow-up, there is a risk of overestimating sensitivity and NPV because of missed PE cases. The meticulousness of follow-up varied between studies. Last, it was not clear in all studies whether cases with an index test positive for PE were followed up, and two studies explicitly did not follow up these cases. In studies in which all cases did undergo the reference test, the aim most often was to identify PE during follow-up, not to exclude it. We dealt with this partial verification by reporting only on sensitivity and NPV. These measures are useful in providing answers to the clinically most relevant question of whether it is safe to withhold treatment after a negative and inconclusive scan. However there is always a trade-off between the sensitivity and specificity; therefore, important information is missing from the included studies.

Although a retrospective design is not considered to carry high risk of bias in the QUADAS-2 assessment, retrospective studies are potentially at risk for selection bias and for incomplete or inaccurate data collection. For instance, if inconclusive scans are at risk of not being registered in the radiology information system, which serves as the reference standard, the percentage of inconclusive scans might be underestimated.

Applicability of findings to the review question

In general, one should apply the results of this review with caution because of the low quality and high degree of heterogeneity of the included studies. Owing to above described limitations of conventional test accuracy measures when a third diagnostic category is included, we regard inconclusives as negative in both NPV and sensitivity. This means that reported accuracy measures apply to negative and inconclusive results as a group, whereas clinicians might be particularly interested in the risk of PE after a negative index test. NPV and sensitivity should be regarded as a conservative best guess.

The above described limitations also translate into uncertainty in the estimate of PE prevalence. For this estimation, we regarded all patients with a positive index test as having had PE, which assumes a 100% positive predictive value. We regarded treatment for PE after an inconclusive test as a positive index test, which also might lead to an overestimation of prevalence.

Overall, study samples appeared representative of the target population. We included patients with pregnancies in all trimesters. When described, the presenting symptoms appeared representative of the population of pregnant patients with suspicion of PE. We included both secondary and tertiary care inpatients and outpatients.

Non-systematic selection between available index tests might hamper the generalisability of results. This occurred in four studies (Bajc 2015; Hullah 2011; Ridge 2009; Shahir 2010). This study design confers the risk that a certain patient group might be selected for one index test, and another for the second index test.

In one of the CTPA studies, patients underwent single-detector or multi-detector computed tomography (CT) (Nijkeuter 2013). The other studies used multi-detector CT scanners for all patients. Two of the scintigraphy studies used only perfusion scintigraphy after chest X-ray (Scarsbrook 2007; Shahir 2010). This is a common diagnostic strategy in clinical practice. Investigators in the other scintigraphy studies performed ventilation scanning in all or a proportion of the study sample, although this was unclear in one study (Cutts 2014). One study used SPECT lung scintigraphy (Bajc 2015), which might be viewed as a separate technique. Four studies employed multiple scanning protocols (Nijkeuter 2013; Ridge 2009; Scarsbrook 2007; Shahir 2010). Results should be taken to as applying to lung scintigraphy or CTPA when the same technique is used.

When possible, we excluded from analysis any second imaging with the same index test. Review results thus apply to first-time scanning. If a different index test was performed before, then the result could still be included (e.g. lung scintigraphy after an inconclusive CTPA). Review results therefore apply to the index test, regardless of prior testing, except when the same index test was used.When reported, we included the original assessment of the index test. Four studies reported only re-assessments for study purposes (Browne 2014; Chan 2002; Cutts 2014; Litmanovich 2009). Results of these studies are less applicable to the clinical situation.

AUTHORS' CONCLUSIONS

Implications for practice

We could not provide summary estimates of test accuracy and therefore can give no direct answer to the research question. In general, both CTPA and lung scintigraphy seem appropriate for exclusion of PE in pregnant women. The median NPV of both CTPA and lung scintigraphy was 100%, despite the conservative approach. An NPV of 100% implies that after a negative or inconclusive index test, the chance that the patient does not have a PE is 100%. This percentage is a median, and a few

(possible) PE cases were in fact missed. Nonetheless, results suggest that the chance of PE after a negative or inconclusive CTPA or lung scintigraphy is marginal. Especially a negative CTPA or lung scintigraphy appears to safely exclude PE during pregnancy.

Median sensitivity of CTPA was 83%. Median sensitivity of lung scintigraphy was 100%. For both, sensitivity ranged from 0% to 100%. This wide range reflects both heterogeneity in the studies and the low prevalence of PE. Most cases of PE that were "missed" involved patients with inconclusive scans. One patient in all CTPA studies had a PE at follow-up after a negative CTPA (Shahir 2010). No patients with a negative lung scintigraphy had a PE at follow-up. In accordance with the high NPV, this implies that both CTPA and lung scintigraphy are appropriate for excluding PE in pregnant women. Given heterogeneity in these studies, the low quality of evidence and necessary assumptions for the analysis, one should adopt this conclusion cautiously.

The difference in median sensitivity between CTPA and lung scintigraphy does not justify the conclusion that lung scintigraphy is more sensitive than CTPA. First, the statistic used is the median, which impairs comparison in this small sample size. Second, no direct comparison was made or statistically tested. Moreover, the lung scintigraphy studies applied different imaging techniques. It is as yet unclear which test has the highest accuracy for PE in pregnancy. Policy makers need to consider other factors, including radiation exposure to mother and foetus, adverse events, patient burden, costs and availability. These aspects fall outside the scope of this review.

Also of clinical importance is the risk of missed PE after an inconclusive scan, which can be expected to be higher than after a negative scan. This review identified very few patients with PE after an inconclusive index test. However, the absolute numbers of inconclusive results were low for both CTPA and lung scintigraphy. Also cases with high clinical suspicion were treated despite the absence of a positive scan. It cannot be concluded on the basis of evidence presented in this review that withholding treatment after an inconclusive test is safe. In clinical practice, most patients go on for further imaging.

Last, this review found that in a population of pregnant women suspected of PE, the median frequency of PE was 3.3%, which is consistent with findings from a recent systematic review that described a rate of diagnosis of venous thromboembolism (VTE) in pregnant women with suspicion of PE at 4.1% (Kline 2014). The pretest probability in this setting is thus low. However, given the potential lethality of PE and the poor performance of clinical judgement in pregnant patients, diagnostic testing should not be withheld.

Implications for research

The finding of a high NPV for both CTPA and lung scintigraphy in this review is based on studies with high risk of bias and concerns about applicability. High-quality direct comparisons of these imaging tests are needed. MRA should be included in this comparison, as this imaging technique avoids the use of radiation. The invasive nature of pulmonary angiography renders it unethical for use as a reference standard. Clinical follow-up is poorly able to identifying false positives; therefore studies with this reference standard can be used only to make inferences about sensitivity and NPV. Sensitivity and NPV are clinically the most important, for safe exclusion of PE in a population in which the a priori risk is already low.

These imaging tests might alternatively be investigated in a randomised controlled trial (RCT). The advantage of this approach is that effects of different tests on downstream patient outcomes could be compared, and this is the relevant endpoint. Outcome measures could include missed PE and bleeding complications associated with treatment decisions based on test results. However, long-term radiation effects cannot be investigated in this way. This is a relevant factor in choosing among tests for pregnant patients. Furthermore, aspects such as costs and availability are not taken into account. A reliable estimate of sensitivity and NPV acquired through rigorous diagnostic test accuracy studies is therefore more valuable. Such estimates can be used for decision making in different clinical settings, not merely in the setting tested by the RCT. Accuracy estimates are robust against the dynamic of changing treatment strategies, modelling assumptions about radiation effects, costs, availability and other contingencies.

Ideally, therefore, CTPA, lung scintigraphy and MRA would be investigated in a large prospective diagnostic test accuracy study with meticulous clinical follow-up as a reference standard. Investigating SPECT instead of planar lung scintigraphy should be considered because of the high level of performance of SPECT in one of the included studies (Bajc 2015), and on the basis of recommendations provided by European guidelines (Bajc 2009; Konstantinides 2014). Availability is an issue with SPECT. Future studies should take into account the Standards for Reporting of Diagnostic Accuracy (STARD) criteria (Bossuyt 2015) and might apply randomisation to ensure comparability. Researchers should clearly define the different test result categories along with prespecified cut-offs. Clear reporting of trimesters is important in that both diagnostic performance and safety profiles of these modalities might vary with the stage of pregnancy. Outcome measures should include sensitivity and NPV for both a negative test result and an inconclusive test result.

ACKNOWLEDGEMENTS

![Cochrane Library - Trusted evidence. Informed decisions. Better health.]

Cochrane Database of Systematic Reviews

REFERENCES

References to studies included in this review

Bajc 2015 *{published data only}*

Bajc M, Olsson B, Gottsater A, Hindorf C, Jogi J. V/P SPECT as a diagnostic tool for pregnant women with suspected pulmonary embolism. *European Journal of Nuclear Medicine and Molecular Imaging* 2015;**42**(8):1325-30.

Bourjeily 2012 *{published data only}*

Bourjeily G, Khalil H, Raker C, Martin S, Auger P, Chalhoub M, et al. Outcomes of negative multidetector computed tomography with pulmonary angiography in pregnant women suspected of pulmonary embolism. *Lung* 2012;**190**(1):105-11.

Browne 2014 *{published data only}*

Browne AM, Cronin CG, NiMhuircheartaigh J, Donagh C, Morrison JJ, Lohan DG, et al. Evaluation of imaging quality of pulmonary 64-MDCT angiography in pregnancy and puerperium. *American Journal of Roentgenology* 2014;**202**(1):60-4.

Chan 2002 *{published data only}*

Chan WS, Ray JG, Murray S, Coady GE, Coates G, Ginsberg JS. Suspected pulmonary embolism in pregnancy: clinical presentation, results of lung scanning, and subsequent maternal and pediatric outcomes. *Archives of Internal Medicine* 2002;**162**(10):1170-5.

Cutts 2014 *{published data only}*

Cutts BA, Tran HA, Merriman E, Nandurkar D, Soo G, DasGupta D, et al. The utility of the Wells clinical prediction model and ventilation-perfusion scanning for pulmonary embolism diagnosis in pregnancy. *Blood Coagulation and Fibrinolysis* 2014;**25**(4):375-8.

Hullah 2011 *{published data only}*

Hullah P, Harris B, Hibbert M. A case review study of investigating for pulmonary embolism in pregnancy. *Respirology* 2011;**16 (Suppl 1)**:67.

Litmanovich 2009 *{published data only}*

Litmanovich D, Boiselle PM, Bankier AA, Kataoka ML, Pianykh O, Raptopoulos V. Dose reduction in computed tomographic angiography of pregnant patients with suspected acute pulmonary embolism. *Journal of Computer Assisted Tomography* 2009;**33**(6):961-6.

Nijkeuter 2013 *{published data only}*

Nijkeuter M, Tan M, Middeldorp S, Kroft LJM, Beenen L, Huisman MV. Safety of ruling out pulmonary embolism (PE) in pregnancy by computed tomography pulmonary angiography (CTPA). *Journal of Thrombosis and Haemostasis* 2013;**11(Suppl 2)**:130.

Ridge 2009 *{published data only}*

Ridge CA, McDermott S, Freyne BJ, Brennan DJ, Collins CD, Skehan SJ. Pulmonary embolism in pregnancy: comparison of pulmonary CT angiography and lung scintigraphy. *American Journal of Roentgenology* 2009;**193**(5):1223-7.

Scarsbrook 2007 *{published data only}*

Scarsbrook AF, Bradley KM, Gleeson FV. Perfusion scintigraphy: diagnostic utility in pregnant women with suspected pulmonary embolic disease. *European Radiology* 2007;**17**(10):2554-60.

Shahir 2010 *{published data only}*

Shahir K, Goodman LR, Tali A, Thorsen KM, Hellman RS. Pulmonary embolism in pregnancy: CT pulmonary angiography versus perfusion scanning. *American Journal of Roentgenology* 2010;**195**(3):W214-20.

References to studies excluded from this review

Balan 1997 *{published data only}*

Balan KK, Critchley M, Vedavathy KK, Smith ML, Vinjamuri S. The value of ventilation-perfusion imaging in pregnancy. *British Journal of Radiology* 1997;**70**(832):338-40.

Cahill 2009 *{published data only}*

Cahill AG, Stout MJ, Macones GA, Bhalla S. Diagnosing pulmonary embolism in pregnancy using computed-tomographic angiography or ventilation-perfusion. *Obstetrics and Gynecology* 2009;**114**(1):124-9.

Herédia 2012 *{published data only}*

Herédia V, Altun E, Ramalho M, de Campos R, Azevedo R, Pamuklar E, et al. MRI of pregnant patients for suspected pulmonary embolism: steady-state free precession vs postgadolinium 3D-GRE. *Acta Medica Portuguesa* 2012;**25**(6):359-67.

Revel 2011 *{published data only}*

Revel M, Cohen S, Sanchez O, Collignon M, Thiam R, Redheuil A, et al. Pulmonary embolism during pregnancy: diagnosis with lung scintigraphy or CT angiography?. *Radiology* 2011;**258**(2):590-8.

Ridge 2011 *{published data only}*

Ridge CA, Mhuircheartaigh JN, Dodd JD, Skehan SJ. Pulmonary CT angiography protocol adapted to the hemodynamic effects of pregnancy. *American Journal of Roentgenology* 2011;**197**(5):1058-63.

Soler 2011 *{published data only}*

Soler S, Khalil H, Raker C, Healey T, Curran P, Bourjeily G. Untreated subsegmental pulmonary embolism (SSPE) in high risk women: incidence and outcomes. *Chest* 2011;**140**(4 Meeting Abstracts):449A.

References to studies awaiting assessment

Vin 2009 *{published data only}*

Vin F. Venous disease among pregnant women: clinical aspects and therapeutic management. *Reproduction Humaine et Hormones* 2009;**22**(SPEC. ISS):3-7.

Additional references

Bajc 2009

Bajc M, Neilly J B, Miniati M, Schuemichen C, Meignan M, Jonson B. EANM guidelines for ventilation/perfusion scintigraphy: Part 1. Pulmonary imaging with ventilation/perfusion single photon emission tomography. *European Journal of Nuclear Medicine and Molecular Imaging* 2009;**36**(8):1356-70.

Bank 2003

Bank I, Libourel EJ, Middeldorp S, Van der Meer J, Buller HR. High rate of skin complications due to low-molecular-weight heparins in pregnant women. *Journal of Thrombosis and Haemostasis* 2003;**1**(4):859-61.

Bates 2012

Bates SM, Greer IA, Middeldorp S, Veenstra DL, Prabulos A, Vandvik PO. VTE, thrombophilia, antithrombotic therapy, and pregnancy: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012;**141**(2 Suppl):e691S-736S.

Bipat 2007

Bipat S, Zwinderman AH, Bossuyt PM, Stoker J. Multivariate random-effects approach: for meta-analysis of cancer staging studies. *Academic Radiology* 2007;**14**(8):974-84.

Bossuyt 2015

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ (Clinical research ed.)* 2015;**351**:h5527. [PUBMED: 26511519]

Bounameaux 2010

Bounameaux H, Perrier A, Righini M. Diagnosis of venous thromboembolism: an update. *Vascular Medicine* 2010;**15**(5):399-406.

Bourjeily 2010

Bourjeily G, Chalhoub M, Phornphutkul C, Alleyne TC, Woodfield CA, Chen KK. Neonatal thyroid function: effect of a single exposure to iodinated contrast medium in utero. *Radiology* 2010;**256**(3):744-50.

Bourjeily 2010a

Bourjeily G, Paidas M, Khalil H, Rosene-Montella K, Rodger M. Pulmonary embolism in pregnancy. *Lancet* 2010;**375**(9713):500-12.

Brown 2010

Brown HL, Hiett AK. Deep vein thrombosis and pulmonary embolism in pregnancy: diagnosis, complications, and management. *Clinical Obstetrics and Gynecology* 2010;**53**(2):345-59.

Calder 2005

Calder KK, Herbert M, Henderson SO. The mortality of untreated pulmonary embolism in emergency department patients. *Annals of Emergency Medicine* 2005;**45**(3):302-10.

Chabloz 2001

Chabloz P, Reber G, Boehlen F, Hohlfeld P, De Moerloose P. TAFI antigen and D-dimer levels during normal pregnancy and at delivery. *British Journal of Haematology* 2001;**115**(1):150-2.

Chang 2003

Chang J, Elam-Evans LD, Berg CJ, Herndon J, Flowers L, Seed KA, et al. Pregnancy-related mortality surveillance - United States, 1991-1999. *MMWR Surveillance Summaries* 2003;**52**(SS02):1-8.

Chen 2008

Chen MM, Coakley FV, Kaimal A, Laros RK Jr. Guidelines for computed tomography and magnetic resonance imaging use during pregnancy and lactation. *Obstetrics and Gynecology* 2008;**112**(2 Pt 1):333-40.

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.

Doshi 2008

Doshi SK, Negus IS, Oduko JM. Fetal radiation dose from CT pulmonary angiography in late pregnancy: a phantom study. *British Journal of Radiology* 2008;**81**(968):653-8.

Duran-Mendicuti 2011

Duran-Mendicuti A, Sodickson A. Imaging evaluation of the pregnant patient with suspected pulmonary embolism. *International Journal of Obstetric Anesthesia* 2011;**20**(1):51-9.

Einstein 2007

Einstein AJ, Henzlova MJ, Rajagopalan S. Estimating risk of cancer associated with radiation exposure from 64-slice computed tomography coronary angiography. *JAMA* 2007;**298**(3):317-23. [PUBMED: 17635892]

Greer 2005

Greer IA, Nelson-Piercy C. Low-molecular-weight heparins for thromboprophylaxis and treatment of venous thromboembolism in pregnancy: a systematic review of safety and efficacy. *Blood* 2005;**106**(2):401-7.

Hendee 1986

Hendee WR, Edwards FM. ALARA and an integrated approach to radiation protection. *Seminars in Nuclear Medicine* 1986;**16**(2):142-50.

Higgins 2009

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [updated September 2009]. The Cochrane Collaboration, 2009. www.cochrane-handbook.org.

Hurwitz 2006

Hurwitz LM, Yoshizumi TT, Reiman RE, Paulson EK, Frush DP, Nguyen GT, et al. Radiation dose to the female breast from 16-MDCT body protocols. *American Journal of Roentgenology* 2006;**186**(6):1718-22.

Hurwitz 2006a

Hurwitz LM, Yoshizumi T, Reiman RE, Goodman PC, Paulson EK, Frush DP, et al. Radiation dose to the fetus from body MDCT during early gestation. *American Journal of Roentgenology* 2006;**186**(3):871-6.

Kanal 1992

Kanal E, Shellock FG. Policies, guidelines, and recommendations for MR imaging safety and patient management. SMRI Safety Committee. *Journal of Magnetic Resonance Imaging* 1992;**2**(2):247-8.

Kline 2014

Kline JA, Richardson DM, Than MP, Penaloza A, Roy P. Systematic review and meta-analysis of pregnant patients investigated for suspected pulmonary embolism in the emergency department. *Academic Emergency Medicine* 2014;**21**(9):949-59.

Kluge 2004

Kluge A, Muller C, Hansel J, Gerriets T, Bachmann G. Real-time MR with TrueFISP for the detection of acute pulmonary embolism: initial clinical experience. *European Radiology* 2004;**14**(4):709-18.

Kobayashi 2008

Kobayashi T, Nakabayashi M, Ishikawa M, Adachi T, Kobashi G, Maeda M, et al. Pulmonary thromboembolism in obstetrics and gynecology increased by 6.5-fold over the past decade in Japan. *Circulation Journal* 2008;**72**(5):753-6.

Konstantinides 2014

Konstantinides SV, Torbicki A, Agnelli G, Danchin N, Fitzmaurice D, Galiè N, et al. 2014 ESC guidelines on the diagnosis and management of acute pulmonary embolism. *European Heart Journal* 2014;**35**(43):3033-69. [PUBMED: 25173341]

Lambert 2002

Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *Journal of Clinical Epidemiology* 2002;**55**(1):86-94.

Leeflang 2006

Leeflang MM, Scholten RJ, Rutjes AW, Reitsma JB, Bossuyt PM. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *Journal of Clinical Epidemiology* 2006;**59**(3):234-40.

Leeflang 2012

Leeflang MM, Deeks JJ, Rutjes AW, Reitsma JB, Bossuyt PM. Bivariate meta-analysis of predictive values of diagnostic tests can be an alternative to bivariate meta-analysis of sensitivity and specificity. *Journal of Clinical Epidemiology* 2012;**65**(10):1088-97.

Lewis 2007

Lewis G (editor). The confidential enquiry into maternal and childHealth (CEMACH). Saving mothers' lives: reviewing maternal deaths to make motherhood safer - 2003-2005. The seventh report on confidential enquiries into maternal deaths in the United Kingdom. London: CEMACH 2007.

Matthews 2006

Matthews S. Short communication: imaging pulmonary embolism in pregnancy: what is the most appropriate imaging protocol?. *British Journal of Radiology* 2006;**79**(941):441-4.

Meng 2015

Meng K, Hu X, Peng X, Zhang Z. Incidence of venous thromboembolism during pregnancy and the puerperium: a systematic review and meta-analysis. *Journal of Maternal-fetal & Neonatal Medicine* 2015;**28**(3):245-53.

Middeldorp 2011

Middeldorp S. How I treat pregnancy-related venous thromboembolism. *Blood* 2011;**118**(20):5394-400.

Morton 2004

Morton SC, Adams JL, Suttorp MJ, Shekelle PG. Meta-regression approaches: what, why, when, and how? Technical review 8 (prepared by Southern California-RAND Evidence-Based Practice Center, under Contract No 290-97-0001). Rockville, MD: Agency for Healthcare Research & Quality 2004.

Mos 2009

Mos IC, Klok FA, Kroft LJ, De Roos A, Dekkers OM, Huisman MV. Safety of ruling out acute pulmonary embolism by normal computed tomography pulmonary angiography in patients with an indication for computed tomography: systematic review and meta-analysis. *Journal of Thrombosis and Haemostasis* 2009;**7**(9):1491-8.

Nijkeuter 2004

Nijkeuter M, Geleijns J, De Roos A, Meinders AE, Huisman MV. Diagnosing pulmonary embolism in pregnancy: rationalizing fetal radiation exposure in radiological procedures. *Journal of Thrombosis and Haemostasis* 2004;**2**(10):1857-8.

Nijkeuter 2006

Nijkeuter M, Ginsberg JS, Huisman MV. Diagnosis of deep vein thrombosis and pulmonary embolism in pregnancy: a systematic review. *Journal of Thrombosis and Haemostasis* 2006;**4**(3):496-500.

Oudkerk 2002

Oudkerk M, van Beek EJ, Wielopolski P, van Ooijen PM, Brouwers-Kuyper EM, Bongaerts AH, et al. Comparison of contrast-enhanced magnetic resonance angiography and conventional pulmonary angiography for the diagnosis of pulmonary embolism: a prospective study. *Lancet* 2002;**359**(9318):1643-7.

PIOPED 1990

PIOPED. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (PIOPED). The PIOPED Investigators. *JAMA* 1990;**263**(20):2753-9.

Pomp 2008

Pomp ER, Lenselink AM, Rosendaal FR, Doggen CJ. Pregnancy, the postpartum period and prothrombotic defects: risk of venous thrombosis in the MEGA study. *Journal of Thrombosis and Haemostasis* 2008;**6**(4):632-7.

Quiroz 2005

Quiroz R, Kucher N, Zou KH, Kipfmueller F, Costello P, Goldhaber SZ, et al. Clinical validity of a negative computed tomography scan in patients with suspected pulmonary embolism: a systematic review. *JAMA* 2005;**293**(16):2012-7.

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.

Ritchie 2007

Ritchie G, Glanville J, Lefebvre C. Do published search filters to identify diagnostic test accuracy studies perform adequately?. *Health Information and Libraries Journal* 2007;**24**(3):188-92.

Rodger 2010

Rodger M. Evidence base for the management of venous thromboembolism in pregnancy. Hematology American Society of Hematology Education Program Book. 2010, issue 1:173-80.

Rosenberg 2007

Rosenberg VA, Lockwood CJ. Thromboembolism in pregnancy. *Obstetrics and Gynecology Clinics of North America* 2007;**34**(3):481-500.

Somarouthu 2010

Somarouthu B, Abbara S, Kalva SP. Diagnosing deep vein thrombosis. *Postgraduate Medicine* 2010;**122**(2):66-73.

Sostman 2008

Sostman HD, Stein PD, Gottschalk A, Matta F, Hull R, Goodman L. Acute pulmonary embolism: sensitivity and specificity of ventilation-perfusion scintigraphy in PIOPED II study. *Radiology* 2008;**246**(3):941-6.

Streffer 2003

Streffer C, Shore R, Konermann G, Meadows A, Uma Devi P, Preston Withers J, et al. Biological effects after prenatal irradiation (embryo and fetus). A report of the International Commission on Radiological Protection. *Annals of the ICRP* 2003;**33**(1-2):5-206.

Tan 2011

Tan M, Huisman MV. The diagnostic management of acute venous thromboembolism during pregnancy: recent advancements and unresolved issues. *Thrombosis Research* 2011;**127 Suppl 3**:S13-6.

Thompson 2002

Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted?. *Statistics in Medicine* 2002;**21**(11):1559-73.

Torbicki 2014

Torbicki A, Perrier A, Konstantinides S, Agnelli G, Galie N, Pruszczyk P, et al. 2014 ESC Guidelines on the diagnosis and management of acute pulmonary embolism. *European Heart Journal* 2014;**35**(43):3033-73.

Tremblay 2012

Tremblay E, Thérasse E, Thomassin-Naggara I, Trop I. Quality initiatives: guidelines for use of medical imaging during pregnancy and lactation. *Radiographics* 2012;**32**(3):897-911.

Webb 2005

Webb JA, Thomsen HS, Morcos SK, Members of Contrast Media Safety Committee of European Society of Urogenital Radiology (ESUR). The use of iodinated and gadolinium contrast media during pregnancy and lactation. *European Radiology* 2005;**15**(6):1234-40.

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.

Winer-Muram 2002

Winer-Muram HT, Boone JM, Brown HL, Jennings SG, Mabie WC, Lombardo GT. Pulmonary embolism in pregnant patients: fetal radiation dose with helical CT. *Radiology* 2002;**224**(2):487-92.

References to other published versions of this review

de Jong 2014

de Jong PG, van Mens TE, Leeflang MMG, Middeldorp S, Nijkeuter M. Imaging for the exclusion of pulmonary embolism in pregnancy. *Cochrane Database of Systematic Reviews* 2014, Issue 3. [DOI: 10.1002/14651858.CD011053]

C H A R A C T E R I S T I C S   O F   S T U D I E S

Characteristics of included studies *[ordered by study ID]*

Bajc 2015

Study characteristics	
Patient sampling	Aim of the study: to assess the prevalence of PE and other lung diseases among pregnant women with suspected PE and to calculate radiation exposure.

Bajc 2015 *(Continued)*

	Type of study: retrospective cohort.
	Enrolled/eligible: 127/188.
	Inclusion period: 2009–2013.
Patient characteristics and setting	Inclusion criteria: pregnant women with suspicion of PE.
	Exclusion criteria: none.
	Setting: Sweden, tertiary care, 83% inpatients, 17% outpatients.
	Age: 30 years, 18–48 (mean, range).
	Gestational age: 24% 1st trimester; 46% 2nd trimester; 30% 3rd trimester.
	Presenting symptoms: 54% dyspnoea, 41% chest pain, 12% cough, 6% calf/thigh swelling, 5% haemoptysis, 2% collapse.
	Suspicion of DVT: 5.5%.
	Prior testing: 15/127 had chest X-ray, 6 of which were abnormal. 5/127 had compression ultrasonography, 1 of which showed DVT.
	Comorbidity: unclear.
	Anticoagulant therapy within 24 hours before testing: unclear.
	Second presentations included: unclear.
	Re-scan after inconclusive scan included in results: no re-scans.
Index tests	Index test: SPECT.
	Original assessment or re-assessment for study: original.
	Assessed by: not specified.
	Diagnostic criteria: not specified.
	Single or multiple protocols used during study: multiple according to trimester.
	Ventilation scanning performed: in proportion of patients.
	Technical specifications: 1- or 2-day protocol. Perfusion 99mTc-MAA 50 or 120 MBq. Venilation 99mTc-Technegas 30 MBq.
Target condition and reference standard(s)	Target condition: PE.
	Reference standard: Patient files from all hospitals in the region were checked for potential later PE or DVT.
Flow and timing	Duration of follow-up: during the same pregnancy and puerperal period.
	Loss to follow-up: none.
	Criteria for choosing between index tests: unclear, in part according to after hours availability.
Comparative	
Notes	The 61 CTPA scans performed in this study were excluded because they did not meet our inclusion criteria regarding the reference standard.
	Study authors provided additional data.

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Bajc 2015 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		High	

Bourjeily 2012

Study characteristics

Bourjeily 2012 *(Continued)*

Patient sampling	Aim of the study: to evaluate VTE outcomes in patients with negative MDCT-PA during pregnancy.
	Type of study: retrospective cohort.
	Enrolled/eligible: 343/343.
	Inclusion period: April 2004-June 2008.
Patient characteristics and setting	Inclusion criteria: pregnant women with suspicion of PE who underwent MDCT.
	Exclusion criteria: none.
	Setting: USA, tertiary care, inpatients and outpatients.
	Age: 29 ± 6.7 years (mean ± SD).
	Gestational age: unclear.
	Presenting symptoms: dyspnoea 75.6%, pleuritic pain 19.5%, non-pleuritic chest pain 45.6%, cough 18.3%, calf swelling 14.3%, wheezing 5.2%, haemoptysis 0.9%.
	Suspicion of DVT: in at least 14.3%.
	Prior testing: unclear.
	Comorbidity: lung disease n = 84 (asthma 79.5%, prior VTE 14.5%, pulmonary hypertension 0.02%, sarcoidosis 0.01%, lung metastases 0.01%, pneumothorax 0.01%), heart disease n = 25, obstructive sleep apnoea n = 3.
	Anticoagulant therapy within 24 hours before testing: at least 2.1%.
	Second presentations included: yes, 6/349 scans.
	Re-scan after inconclusive scan included in results: no re-scans.
Index tests	Index test: CTPA.
	Original assessment or re-assessment for study: original.
	Assessed by: board-certified radiologist.
	Diagnostic criteria: Study was labelled as ''non-diagnostic'' subjectively if vessel opacification was poor and PE could not be excluded at the segmental level or more proximally.
	Single or multiple protocols used during study: interpretation on film or on a Picture Archiving and Communication System (PACS). Otherwise, the same.
	Technical specifications: 4-MDCT. Collimation 4 × 2.5 mm. Rotation time 0.8 second. Pitch 1.5–1. 120 kV. Auto mAs.
	Further details: Data used for current review reflect scans as unit of analysis. Data with individual patients as unit of analysis could not be extracted.
Target condition and reference standard(s)	Target condition: PE.
	Reference standard: medical record review of own institution and imaging study query at 2 other institutions for evidence of VTE. If no data were available, the patient was considered loss to follow-up.
Flow and timing	Duration of follow-up: 3 months or 6 weeks postpartum (whichever came later).
	Loss to follow-up: 16.
	Criteria for choosing between index tests: NA.

Bourjeily 2012 *(Continued)*

Comparative	
Notes	Study authors provided additional data.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Low	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		High	

Browne 2014

Study characteristics	
Patient sampling	Aim of the study: to quantitatively and qualitatively evaluate pulmonary 64-MDCT angiography image quality during pregnancy and puerperium. Type of study: retrospective cohort. Enrolled/eligible: 70/70. Inclusion period: 3-year period.
Patient characteristics and setting	Inclusion criteria: women with suspicion of PE during pregnancy and puerperium and matched controls. Exclusion criteria: none. Setting: Ireland. Age: 31.1 ± 6.1 years (mean ± SD). Gestational age: 1.4% 1st trimester; 44.3% 2nd trimester; 54.3% 3rd trimester. Presenting symptoms: unclear. Suspicion of DVT: unclear. Prior testing: 5 of the pregnant or puerperium patients had 2 scans, and 1 had 3 scans. Other prior testing was unclear. Comorbidity: unclear. Anticoagulant therapy within 24 hours before testing: unclear. Second presentations included: unclear. Re-scan after inconclusive scan included in results: yes.
Index tests	Index test: CTPA. Original assessment or re-assessment for study: re-assessment. Assessed by: radiologists with 7 and 15 years of experience. Diagnostic criteria: not specified. Single or multiple protocols used during study: single. Technical specifications: 64-MDCT. Pitch 0.9. Rotation time 0.5 second. Collimation 32 × 0.6 mm. 100–120 kVp. Current was adjusted to patient weight.
Target condition and reference standard(s)	Target condition: PE at different levels. Segmental or more proximal PE considered as target condition for current review. Reference standard: checking hospital radiology system for subsequent imaging.
Flow and timing	Duration of follow-up: 6 months. Loss to follow-up: none. Criteria for choosing between index tests: NA.

Browne 2014 *(Continued)*

Comparative	
Notes	Attempted to contact study authors in July 2015. We received no reply.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		High	

Chan 2002

Study characteristics	
Patient sampling	Aim of the study: to study the distribution of lung scan results and the safety of V/Q scanning, as well as the safety of withholding anticoagulation therapy following a normal or non-diagnostic scan in pregnant women.
	Type of study: retrospective cohort.
	Enrolled/eligible: 105/105.
	Inclusion period: January 1990-April 2000.
Patient characteristics and setting	Inclusion criteria: pregnant women with suspicion of PE undergoing perfusion scintigraphy.
	Exclusion criteria: full-dose anticoagulation therapy.
	Setting: Canada, 53% secondary and 47% tertiary care.
	Age: 32 (17-41) years.
	Gestational age: 9.9% 1st trimester; 42.5% 2nd trimester; 47.9% 3rd trimester.
	Presenting symptoms: 62% dyspnoea, 46% pleuritic pain, 19% non-pleuritic chest pain.
	Suspicion of DVT: unclear.
	Prior testing: time relation of testing unclear. 50% had chest radiograph, 40% of which were abnormal. 61% had leg ultrasonography or impedance plethysmography. No patient had a DVT.
	Comorbidity: unclear.
	Anticoagulant therapy within 24 hours before testing: none.
	Second presentations included: no.
	Re-scan after inconclusive scan included in results: no re-scans.
Index tests	Index test: lung scintigraphy.
	Original assessment or re-assessment for study: re-assessment.
	Assessed by: 2 experts.
	Diagnostic criteria: high probability: > 75% subsegmental or greater perfusion defects with normal ventilation.
	Single or multiple protocols used during study: multiple.
	Ventilation scintigraphy performed: yes.
	Technical specifications: perfusion 99mTc-MAA. Venilation 99mTc methylene diphosphonate aerosol or 99mTc sulphur colloid.
Target condition and reference standard(s)	Target condition: pulmonary embolism.
	Reference standard: clinical follow-up. In part by telephone, otherwise through family physician and review of medical records.
	Further details: 2 patients had pulmonary angiography, 1 of which was positive. This patient had a non-diagnostic index test but was excluded because of therapeutic anticoagulation.
Flow and timing	Duration of follow-up: for telephone contact 20.6 months, 0.5-108.0 (mean, range). Unclear duration for GP contact and record review.

Chan 2002 *(Continued)*

| | Loss to follow-up: none. |
| | Criteria for choosing between index tests: NA. |

Comparative	
Notes	Study authors were contacted but could not provide additional data.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	**High**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	No		

Chan 2002 *(Continued)*

High

Cutts 2014

Study characteristics	
Patient sampling	Aim of the study: to determine if clinical prediction models have potential diagnostic validity.
	Type of study: retrospective cohort.
	Enrolled/eligible: 183/216.
	Inclusion period: 2007-2010.
Patient characteristics and setting	Inclusion criteria: pregnant women who underwent imaging for PE.
	Exclusion criteria: unclear.
	Setting: 2 tertiary centres in Australia and the United Kingdom.
	Age: 30 years; 18-44 (median; range).
	Gestational age: 12% 1st trimester; 33% 2nd trimester; 55% 3rd trimester.
	Presenting symptoms: unclear.
	Suspicion of DVT: unclear.
	Prior testing: modified Wells determined retrospectively: 58% likely and 42% unlikely. D-dimer was performed in 51 women: 48 were positive.
	Comorbidity: unclear.
	Anticoagulant therapy within 24 hours before testing: unclear.
	Second presentations included: NA.
	Re-scan after inconclusive scan included in results: no re-scans.
Index tests	Index test: lung scintigraphy.
	Original assessment or re-assessment for study: re-assessment.
	Assessed by: 2 experienced radiologists.
	Diagnostic criteria: not specified.
	Single or multiple protocols used during study: not specified.
	Ventilation scintigraphy performed: not specified.
Target condition and reference standard(s)	Target condition: PE.
	Reference standard: review of medical record.
Flow and timing	Duration of follow-up: unclear, but through postpartum period.
	Loss to follow-up: none.
	Criteria for choosing between index tests: NA.

Cutts 2014 *(Continued)*

	Comparative
Notes	Study authors provided additional data.
	Portion of cohort used for earlier abstract publication, "The utility of the wells clinical prediction model and ventilation-perfusion scanning for pulmonary embolism in pregnancy", published in *Journal of Thrombosis and Haemostasis* in 2011.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		High	

Hullah 2011

Study characteristics	
Patient sampling	Aim of the study: to determine the incidence of PE in pregnant patients investigated for this condition.
	Type of study: retrospective cohort.
	Enrolled/eligible: 49/49.
	Inclusion period: January 2005-January 2010.
Patient characteristics and setting	Inclusion criteria: pregnant women with suspicion of PE. Retrospectively identified by positive bHCG collected at the time of the investigation.
	Exclusion criteria: none.
	Setting: Australia, tertiary care hospital.
	Age: 32.6 ± 5.5 years (mean ± SD).
	Gestational age: 4% 1st trimester; 53% 2nd trimester; 39% 3rd trimester; 4% unknown.
	Presenting symptoms: unclear.
	Suspicion of DVT: unclear.
	Prior testing: 14 positive D-dimer and 1 negative in 49 patients. 27 of 44 patients had a chest X-ray. 18 of 42 patients had compression ultrasonography, which revealed no cases of DVT. 1 patient in the scintigraphy group had a prior CTPA.
	Comorbidity: unclear.
	Anticoagulant therapy within 24 hours before testing: At least 7/49 received clexane.
	Second presentations included: NA.
	Re-scan after inconclusive scan included in results: yes, unclear number of re-scans.
Index tests	Index test: lung scintigraphy.
	Original assessment or re-assessment for study: not specified.
	Assessed by: not specified.
	Diagnostic criteria: not specified.
	Single or multiple protocols used during study: not specified.
	Ventilation scintigraphy performed: in all.
	Technical specifications: perfusion 99mTc-MAA 100 MBq. Ventilation Technegas 20-25 MBq.
Target condition and reference standard(s)	Target condition: PE.
	Reference standard: medical record review for patients with initial positive scan and for patients who re-presented to the hospital.
Flow and timing	Duration of follow-up: until delivery.
	Loss to follow-up: none.
	Criteria for choosing between index tests: clinician and patient preference.

Hullah 2011 *(Continued)*

	Comparative
Notes	We excluded the 21 CTPA scans performed in this study for not meeting our inclusion criteria of a minimum sample size of 25.
	The publication was in abstract form.
	Study authors provided additional data.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Unclear**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	**High**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	No		

Hullah 2011 *(Continued)*

Were all patients included in the analysis?	Yes

	High

Litmanovich 2009

Study characteristics	

Patient sampling	Aim of the study: to assess the effect of this reduced-dose pulmonary CTA protocol on radiation dose and image quality in pregnant patients.
	Type of study: retrospective cohort.
	Enrolled/eligible: 26/unclear.
	Inclusion period: July 2006 to July 2007.
Patient characteristics and setting	Inclusion criteria: pregnant women with a suspicion of PE undergoing CTA.
	Exclusion criteria: known allergy to contrast material, renal insufficiency, suspected hyperparathyroidism.
	Setting: USA. Emergency department.
	Age: 29 ± 5 years (mean ± SD)
	Gestational age: 19 ± 10 (mean ± SD)
	Presenting symptoms: unclear.
	Suspicion of DVT: unclear.
	Prior testing: D-dimer in proportion of women, results unclear. Compression ultrasonography in 18, negative in 18.
	Comorbidity: unclear.
	Anticoagulant therapy within 24 hours before testing: unclear.
	Second presentations included: NA.
	Re-scan after inconclusive scan included in results: no re-scans.
Index tests	Index test: CTPA.
	Original assessment or re-assessment for study: re-assessment.
	Assessed by: radiologists with 7 and 5 years of experience.
	Diagnostic criteria: not specified.
	Single or multiple protocols used during study: single.
	Technical specifications: 64-row multi-detector CT. Collimation 0.625 × 64 mm. Pitch 0.984. Rotation time 0.35 second. 100 kVp. 100 mAs.
Target condition and reference standard(s)	Target condition: segmental PE.
	Reference standard: hospital information system checked for rehospitalisation and PE or DVT diagnosis after index test.

Litmanovich 2009 *(Continued)*

Flow and timing	Duration of follow-up: 18 ± 5 months (mean ± SD).
	Loss to follow-up: none.
	Criteria for choosing between index tests: NA.

Comparative	

Notes	We attempted to contact study authors by email in July 2015 and received no reply.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

![Cochrane Library logo] Cochrane Trusted evidence.
Library Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Litmanovich 2009 *(Continued)*

	High

Nijkeuter 2013

Study characteristics	
Patient sampling	Aim of the study: to investigate ruling out PE by CTPA in pregnant patients with clinical suspicion of PE. Type of study: retrospective cohort. Enrolled/eligible: 134/141. Inclusion period: February 2004-December 2012.
Patient characteristics and setting	Inclusion criteria: pregnant women with suspicion of PE. Exclusion criteria: treatment with therapeutic dose of heparins, younger than 18 years, allergy to contrast. Setting: the Netherlands, secondary and tertiary care, inpatients and outpatients. Age: 31 years, 27-35 (median, IQR). Gestational age: 8.2% 1st trimester; 40.3% 2nd trimester; 50.0% 3rd trimester; 1.5% unclear. Presenting symptoms: unclear. Suspicion of DVT: none. Prior testing: unclear. Comorbidity: unclear. Anticoagulant therapy within 24 hours before testing: none at therapeutic dose. Prophylactic dose unclear. Second presentations included: unclear. Re-scan after inconclusive scan included in results: no re-scans.
Index tests	Index test: CTPA. Original assessment or re-assessment for study: original assessment. Performed by: local attending radiologist. Diagnostic criteria: if contrast material outlined an intraluminal defect or if a vessel was totally occluded by low-attenuation material on at least 2 adjacent slices. Single or multiple protocols used during study: multiple. Technical specifications: single-detector or multi-detector-row systems. No further specifications.
Target condition and reference standard(s)	Target condition: pulmonary embolism. Reference standard: outpatient visit or telephone interview. Further details: only with negative or inconclusive CTPA.

Nijkeuter 2013 *(Continued)*

	Follow-up consisted of a scheduled outpatient visit or telephone interview at 3 months. Patients were additionally instructed to contact the study centre or their general practitioner in the event of symptoms suggestive of DVT or PE.
Flow and timing	Duration of follow-up: 3 months.
	Loss to follow-up: none.
	Criteria for choosing between index tests: NA.
	Further details: The aim was to assess negative predictive value; therefore, follow-up was performed only for patients without a confirmed PE on initial CTPA.
Comparative	
Notes	Study authors provided additional data.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
		High	High
DOMAIN 4: Flow and Timing			

Nijkeuter 2013 *(Continued)*

Was there an appropriate interval between index test and reference standard?	No
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
	High

Ridge 2009

Study characteristics	
Patient sampling	Aim of the study: to compare diagnostic adequacy of lung scintigraphy vs pulmonary CTPA in pregnant patients with suspected PE and to identify causes of diagnostic inadequacy.
	Type of study: retrospective cohort.
	Enrolled/eligible: 50/50.
	Inclusion period: July 2006-April 2008.
Patient characteristics and setting	Inclusion criteria: pregnant women with suspicion of PE, aged 15-45.
	Exclusion criteria: none.
	Setting: Ireland, maternity hospital emergency department.
	Age: CTPA 32.6 ± 5.6 years, lung scintigraphy 31.8 ± 5.4 years (mean ± SD).
	Gestational age: unclear.
	Presenting symptoms: unclear.
	Suspicion of DVT: none.
	Prior testing: D-dimer was performed in an unknown proportion of patients. All patients in scintigraphy group had normal chest X-ray. 3/25 in CTPA group had abnormal chest X-ray.. Two patients underwent lung scintigraphy after the initial non-diagnostic CTPA, and 1 had lung scintigraphy after 2 non-diagnostic CTPA scans.
	Comorbidity: unknown.
	Anticoagulant therapy within 24 hours before testing: unclear.
	Second presentations included: NA.
	Re-scan after inconclusive scan included in results: yes, 3 re-scans in CTPA group.
Index tests	Index test: CTPA (n = 28), lung scintigraphy (n = 25).
	Original assessment or re-assessment for study: original.
	Assessed by: unclear.
	Diagnostic criteria: not specified.
	For CTPA

Ridge 2009 *(Continued)*

- Single or multiple protocols used during study: single.

- Technical specifications: 64-MDCT scanner. Collimation 0.6 mm. Rotation speed 0.33 second. Pitch 0.9. 120 kV. 200 mAs.

For lung scintigraphy

- Ventilation scintigraphy performed: minority of patients.

- Single or multiple protocols used during study: multiple.

- Technical specifications: perfusion 99mTc-labeled MAA 90 MBq. Venilation 99mTc-carbon particles 485 ± 72 MBq.

Target condition and reference standard(s)	Target condition: PE.
	Reference standard: chart review and imaging review on re-presentation.
Flow and timing	Duration of follow-up: 3 months.
	Loss to follow-up: none.
	Criteria for choosing between index tests: physician preference; if abnormality on initial chest X-ray, CTPA.
Comparative	
Notes	Study authors provided additional data.
	Portion of cohort used for the publication, "Pulmonary CT angiography protocol adapted to the haemodynamic effects of pregnancy", published in *American Journal of Roentgenology* in 2011.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		High	Unclear
DOMAIN 3: Reference Standard			

Ridge 2009 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Low	Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	High	

Scarsbrook 2007

Study characteristics	
Patient sampling	Aim of the study: to determine if the diagnostic utility of perfusion scintigraphy in pregnant patients with suspected PE could be optimised by careful patient selection.
	Type of study: retrospective cohort.
	Enrolled/eligible: 96/96 for lung scintigraphy. Another 9 patients had CTPA.
	Inclusion period: January 2001-December 2005.
Patient characteristics and setting	Inclusion criteria: pregnant women with suspicion of PE who underwent imaging.
	Exclusion criteria: abnormal chest X-ray, asthma or chronic lung disease.
	Setting: United Kingdom.
	Age: 29.8 years (range 16–48).
	Gestational age: 8% 1st trimester; 23% 2nd trimester; 69% 3rd trimester.
	Presenting symptoms: unclear.
	Suspicion of DVT: unclear.
	Prior testing: All had negative chest X-ray.
	Comorbidity: no asthma or chronic lung disease, no other comorbidity reported.
	Anticoagulant therapy within 24 hours before testing: unclear.
	Second presentations included: one.

Scarsbrook 2007 *(Continued)*

	Re-scan after inconclusive scan included in results: no re-scans.
Index tests	Index test: lung scintigraphy.
	Original assessment or re-assessment for study: re-assessment but original assessments were extractable from the article.
	Assessed by: not specified.
	Diagnostic criteria: not specified.
	Ventilation scintigraphy performed: no.
	Single or multiple protocols used during study: multiple.
	Technical specifications: perfusion 99mTc-MAA 40 or 80 MBq.
Target condition and reference standard(s)	Target condition: PE.
	Reference standard: review of institutional obstetrical records for pregnancy outcome and assessment of any imaging during follow-up period.
Flow and timing	Duration of follow-up: reported only for the group with non-diagnostic scan: 744 days; 87–1421 (mean; range).
	Loss to follow-up: none.
	Criteria for choosing between index tests: standardised protocol. Patients with normal chest X-ray and no pulmonary comorbidity underwent scintigraphy; other patients underwent CTPA. Protocol was not followed in 100% of cases.
Comparative	
Notes	Study authors were contacted but could not provide additional data.
	We excluded the 9 patients who had CTPA because this group met the exclusion criterion of a sample size smaller than 25.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		Low	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge	Yes		

Scarsbrook 2007 *(Continued)*

of the results of the reference
standard?

	Low	High
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	**Unclear**
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
	High	

Shahir 2010

Study characteristics	
Patient sampling	Aim of the study: to determine whether CTPA or perfusion scanning has better image quality and to determine the likelihood of subsequent PE after a negative interpretation.
	Type of study: retrospective cohort.
	Enrolled/eligible: 199/unclear.
	Inclusion period: 2000–2007.
Patient characteristics and setting	Inclusion criteria: pregnant patients who underwent CTPA or perfusion scanning for suspected PE.
	Exclusion criteria: DVT diagnosed on ultrasonography.
	Setting: USA, 47% at private hospital, 53% at university hospital.
	Age: 31 years, 18-39 (mean, range).
	Gestational age: 27% 1st trimester; 31% 2nd trimester; 42% 3rd trimester.
	Presenting symptoms: 90% dyspnoea, 65% chest pain, 15% calf/thigh pain, 2% collapse, 1% haemoptysis.
	Suspicion of DVT: unclear.

Shahir 2010 *(Continued)*

Prior testing: abnormal chest X-ray: 8% in CTPA group and 5% in perfusion scan group. The rest had normal chest X-ray. 58% underwent bilateral compression ultrasonography, all negative for DVT. In both CTA and perfusion scan groups, 3 patients underwent the other index test after non-diagnostic scan.

Comorbidity: unclear.

Anticoagulant therapy within 24 hours before testing: 8%.

Second presentations included: unclear.

Re-scan after inconclusive scan included in results: no re-scans.

Index tests	Index test: CTPA (n = 106), lung scintigraphy (n = 99).
	Original assessment or re-assessment for study: All re-assessments were in agreement with the original ones.
	For CTPA
	- Assessed by: 2 radiologists with 4 and 25 years of experience.
	- Diagnostic criteria: filling defect in or failure of opacification of a pulmonary artery.
	- Single or multiple protocols used during study: multiple.
	- Technical specifications: 8-, 16-, or 64-MDCT scanner. Pitch 1. 120–140 kV. 150–300 mAs.
	For lung scintigraphy
	- Assessed by: a fellow in nuclear medicine and in some cases a nuclear medicine specialist (30 years of experience).
	- Diagnostic criteria: threshold of more than 1.5 segmental perfusion defects for high probability.
	- Ventilation scintigraphy performed: no.
	- Technical specifications: 1–1.5 mCi (37–55 MBq) of 99mTc-labeled MAA.
Target condition and reference standard(s)	Target condition: PE.
	Reference standard: clinical follow-up. Information from hospital charts on suspicion, detection or treatment of PE.
Flow and timing	Duration of follow-up: 3 months.
	Loss to follow-up: 2%.
	Criteria for choosing between index tests: unclear.
Comparative	
Notes	Attempted to contact study authors in July 2015. We received no reply.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Shahir 2010 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	**High**

DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**High**

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	**High**

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
		High

99mTc: technetium 99m.
bHCG: beta human chorionic gonadotropin.
CT: computed tomography.
CTA: computed tomography angiography.
CTPA: computed tomography pulmonary angiography.
DVT: deep vein thrombosis.
GP: general practitioner.
IQR: interquartile range.
MAA: macro aggregated albumin.
MDCT: multi-detector computed tomography.
MDCT-PA: multi-detector computed tomography with pulmonary angiography.
NA: not applicable.
PACS: Picture Archiving and Communication System.

PE: pulmonary embolism.
SD: standard deviation.
SPECT: single-photon emission computed tomography.
V/Q: ventilation/perfusion.
VTE: venous thromboembolism.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Balan 1997	Duration of follow-up not described.
Cahill 2009	Discrepancies identified between additional data provided by study authors on the pregnant patient subgroup and data reported in the article, which could not be resolved by study authors.
Herédia 2012	Fewer than 25 pregnant women. Study on MRA (not a reason for exclusion).
Revel 2011	Discrepancies identified between additional data provided by study authors and data reported in the article, which could not be resolved by study authors.
Ridge 2011	Portion of cohort overlaps with included study Ridge 2009. The remaining portion of the cohort, with separate CTPA protocol, consisted of fewer than 25 pregnant women.
Soler 2011	Required data could not be extracted from the published abstract and was not provided by study authors.

CTPA: computed tomography pulmonary angiography.
MRA: magnetic resonance angiography.

Characteristics of studies awaiting classification *[ordered by study ID]*

Vin 2009

Study characteristics	
Patient sampling	
Patient characteristics and setting	
Index tests	
Target condition and reference standard(s)	
Flow and timing	
Comparative	
Notes	Full text could not be obtained.

D A T A

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 Primary analysis CTPA	6	695
2 Primary analysis lung scintigraphy	7	665
3 Sensitivity analysis CTPA	6	695
4 Sensitivity analysis lung scintigraphy	7	665

Test 1. Primary analysis CTPA.

Test 2. Primary analysis lung scintigraphy.

Test 3. Sensitivity analysis CTPA.

Test 4. Sensitivity analysis lung scintigraphy.

ADDITIONAL TABLES

Table 1. QUADAS-2 checklist

	Item plus signalling questions	Criteria for scoring 'yes', 'no' and 'unclear'
1	PATIENT SELECTION Was a consecutive or random sample of patients enrolled? Was a case-control design avoided? Did the study avoid inappropriate exclusions?	We will score this item 'yes' when patients were consecutively or randomly selected; 90% or more were evaluated at the hospital; 5% or less of had received anticoagulant therapy within 24 hours before testing; 30% or less were given a diagnosis of comorbidity such as chronic obstructive pulmonary disease or other pulmonary disease, malignancy or pregnancy complications (preeclampsia, syndrome of haemolysis, elevated liver enzymes and low platelets or eclampsia); and 10% or less had undergone prior testing for this episode of suspected PE. We will score 'no' if one of these criteria was not met.
2	INDEX TEST Were index test results interpreted without knowledge of results of the reference standard?	We will score this item 'yes' in the following cases: if study authors state that the index test interpreter was unaware of the result of the reference test; or if the order of testing was index test before reference test for every patient. Even if clinical follow-up was the reference test, the order of testing has to be stated for the item to be scored 'yes'. We will score the item 'no' for studies in which it is stated that the interpreter of the index test was aware of the result of the reference test. In other cases, we will score this as 'unclear'.

Table 1. QUADAS-2 checklist *(Continued)*

		In cases of studies directly comparing the diagnostic accuracy of 2 index tests against the reference standard, these test results had to be interpreted without knowledge of the results of the comparator index test, and we will score this item similarly to the approach described above.
3	REFERENCE STANDARD Is the reference standard likely to correctly classify the target condition? Were reference standard results interpreted without knowledge of results of the index test?	We considered both PA and clinical follow-up of at least 6 weeks as useful for correct classification of the target condition, the latter only if objective diagnostic tests are used in cases of suspected venous thromboembolism. We will score this item 'yes' if study authors state that reference tests were interpreted without knowledge of results of the index test. Furthermore, in cases of clinical follow-up as a reference standard, any clinical suspicion of venous thrombosis during follow-up needs to be followed by objective diagnostic testing (i.e. CUS or venography for suspicion of DVT, and scintigraphy, CTPA or pulmonary angiography for clinical suspicion of PE). If a patient died during follow-up, we classified death as caused by PE in cases of confirmation by autopsy, in cases of an objective test positive for PE before death or if PE could not be confidently excluded as the cause of death.
4	FLOW AND TIMING Was an appropriate interval between index test(s) and reference standard provided? Did all patients receive a reference standard? Did all patients receive the same reference standard? Were all patients included in the analysis?	With PA, we will consider a time period of less than 24 hours between index and reference tests as short enough to ensure that the target condition did not change between tests, either because of natural progression of the disease or because of therapeutic intervention. For studies using pulmonary angiography as the reference test, we will score this item 'yes' if the time between index and reference tests was less than 24 hours. Similarly, for studies directly comparing diagnostic accuracy of index tests, we will consider a time period of less than 24 hours between index tests and the reference test as short enough. During clinical follow-up, the disease may diminish through natural progression or through intervention. Or the condition may arise during follow-up if it was not present at the time of the index test. Therefore, we will score studies using clinical follow-up 'no' for this item. We will score this item 'no' if less than 90% or a non-random selection of patients underwent the reference test. We will score this item 'no' if less than 90% of patients who had an index test result underwent pulmonary angiography or had clinical follow-up as the reference test.

CTPA: computed tomography pulmonary angiography.
CUS: compression ultrasonography.
DVT: deep vein thrombosis.
PA: pulmonary angiography.
PE: pulmonary embolism.

A P P E N D I C E S

Appendix 1. MEDLINE search strategy

1	Thromboembolism/	20790
2	Venous Thromboembolism/	5497
3	Thrombosis/	57313

(Continued)

4	Venous Thrombosis/	18922
5	Upper Extremity Deep Vein Thrombosis/	194
6	(thromboemboli$ or microthrombus or thrombus or thrombo$ or thrombilic or thrombotic).ti,ab.	298311
7	(DVT or VTE).ti,ab.	11960
8	Pulmonary Embolism/	32316
9	(pulmonary adj embol$).ti,ab.	28329
10	(pulmonary adj thrombo$).ti,ab.	3645
11	(lung adj embol$).ti,ab.	381
12	(lung adj thrombo$).ti,ab.	64
13	PE.ti,ab.	25958
14	or/1-13	368515
15	exp Pregnancy/	732514
16	pregnan$.ti,ab.	390395
17	prepartum.ti,ab.	1747
18	antepartum.ti,ab.	4605
19	prenatal.ti,ab.	71313
20	antenatal.ti,ab.	24738
21	exp Pregnancy Trimesters/	33767
22	trimester?.ti,ab.	42203
23	or/15-22	844362
24	diagnostic imaging/	32966
25	exp Lung/ra, ri, us [Radiography, Radionuclide Imaging, Ultrasonography]	20697
26	Pulmonary Artery/ra, ri, us [Radiography, Radionuclide Imaging, Ultrasonography]	6616
27	Angiography/	54573
28	(pulmonary adj3 angiogr$).ti,ab.	3629
29	(lung adj3 angiogr$).ti,ab.	248
30	magnetic resonance imaging/	303668

(Continued)

31	magnetic resonance angiography/	18622
32	Diffusion Magnetic Resonance Imaging/	12340
33	((magnetic resonance or MR or MRI or NMR) adj5 (angiogra$ or arteriogra$)).ti,ab.	16040
34	MRA.ti,ab.	6223
35	tomography/	9269
36	exp tomography, emission-computed/	83280
37	exp tomography, x-ray/	325764
38	tomography scanners, x-ray computed/	2035
39	((CT or CAT) adj5 (angiogra$ or arteriogra$ or tomograph$)).ti,ab.	77957
40	(comput$ adj3 tomogra$).ti,ab.	207656
41	tomodensitometry.ti,ab.	592
42	(cat adj4 scan$).ti,ab.	1285
43	(ct adj4 scan$).ti,ab.	74762
44	(CTA or CTPA).ti,ab.	6817
45	three dimensional-ct.ti,ab.	1015
46	(MDCT or MSCT).ti,ab.	7493
47	multislice.ti,ab.	5013
48	multi-slice.ti,ab.	1634
49	(multi-row or multirow).ti,ab.	175
50	(single-slice or singleslice).ti,ab.	1145
51	3d-cta.ti,ab.	227
52	exp ultrasonography, doppler/	58611
53	ultrasonography/	63788
54	exp ultrasonography, prenatal/	26369
55	ultrasonics/	20767
56	ultrasound.ti,ab.	166695
57	ultrasonogra$.ti,ab.	82761
58	ultrasonic$.ti,ab.	42986

(Continued)

59	echograph$.ti,ab.	8737
60	(USS or DUS or CDUS or CEUS).ti,ab.	2945
61	(doppler or duplex).ti,ab.	110831
62	sonograph$.ti,ab.	45002
63	sonogram$.ti,ab.	3359
64	(contrast adj4 US).ti,ab.	1506
65	Phlebography/	10804
66	radionuclide imaging/ or perfusion imaging/	25611
67	Ventilation-Perfusion Ratio/	5543
68	(VQ or V?Q or V?P).ti,ab.	34835
69	SPECT.ti,ab.	21984
70	perfusion.ti,ab.	132776
71	scintigraph$.ti,ab.	41522
72	or/24-71	1366959
73	14 and 23 and 72	2303

Appendix 2. Embase search strategy

1	thromboembolism/	58814
2	thrombosis/	118468
3	vein thrombosis/	30660
4	THROMBUS/	19125
5	MICROTHROMBUS/	1060
6	(thrombus or microthrombus or thrombotic or thrombilic or thromboemboli$ or thrombos$).ti,ab.	265666
7	DEEP VEIN THROMBOSIS/	42111
8	deep vein$ thrombo$.ti,ab.	17473
9	deep venous thrombo$.ti,ab.	12470
10	(dvt or VTE).ti,ab.	20840

(Continued)

11	exp lung/	306946
12	pulmonary artery/	42030
13	lung embolism/	71089
14	(pulmonary adj embol$).ti,ab.	43593
15	(pulmonary adj thrombo$).ti,ab.	5203
16	(lung adj embol$).ti,ab.	761
17	(lung adj thrombo$).ti,ab.	112
18	PE.ti,ab.	39080
19	or/1-18	759291
20	exp pregnancy/	669905
21	pregnan$.ti,ab.	536263
22	prepartum.ti,ab.	1840
23	trimester?.ti,ab.	58057
24	antepartum.ti,ab.	6194
25	prenatal.ti,ab.	89611
26	antenatal.ti,ab.	33682
27	or/20-26	895638
28	exp diagnostic imaging/	124191
29	exp nuclear magnetic resonance imaging/	606198
30	angiography/	89251
31	(pulmonary adj3 angiogr$).ti,ab.	5532
32	(lung adj3 angiogr$).ti,ab.	367
33	((magnetic resonance or MR or MRI or NMR) adj5 (angiogra$ or arteriogra$)).ti,ab.	20670
34	MRA.ti,ab.	9552
35	Radiodensitometry/	3954
36	exp tomography/	770782
37	((CT or CAT) adj5 (angiogra$ or arteriogra$ or tomograph$)).ti,ab.	100240
38	(comput$ adj3 tomogra$).ti,ab.	249406

(Continued)

39	tomodensitometry.ti,ab.	697
40	(cat adj4 scan$).ti,ab.	1803
41	(ct adj4 scan$).ti,ab.	114356
42	(CTA or CTPA).ti,ab.	11995
43	three dimensional-ct.ti,ab.	1274
44	(MDCT or MSCT).ti,ab.	11824
45	(multi-slice or multislice or multi-row or multirow).ti,ab.	9184
46	(single-slice or singleslice).ti,ab.	1372
47	3d-cta.ti,ab.	342
48	exp echography/	568375
49	computed tomography scanner/	11875
50	ultrasound/	123540
51	ultrasound.ti,ab.	242300
52	ultrasonogra$.ti,ab.	108294
53	ultrasonic$.ti,ab.	52948
54	echograph$.ti,ab.	11356
55	(USS or DUS or CDUS or CEUS).ti,ab.	5513
56	(doppler or duplex).ti,ab.	146940
57	sonograph$.ti,ab.	58807
58	sonogram$.ti,ab.	4132
59	(contrast adj4 US).ti,ab.	1993
60	exp phlebography/	24842
61	scintiangiography/ or scintiphlebography/	2935
62	computer assisted scintigraphy/	304
63	phlebogra$.ti,ab.	6440
64	lung ventilation perfusion ratio/	6711
65	lung perfusion/	6348
66	(VQ or V?Q or V?P).ti,ab.	46317

(Continued)

67	imag$.ti,ab.	1011000
68	perfusion.ti,ab.	180169
69	scintigraph$.ti,ab.	56484
70	SPECT.ti,ab.	33023
71	lung angiography/	6667
72	or/28-71	2638744
73	19 and 27 and 72	5841

CONTRIBUTIONS OF AUTHORS

The protocol was written by Thijs van Mens and Paulien de Jong, with comments provided by Mathilde Nijkeuter and Saskia Middeldorp. Mariska Leeflang provided support with methodological aspects of the protocol. Thijs van Mens and Luuk Scheres conducted the search and extracted study data. Thijs van Mens wrote the manuscript with input from all review authors and is the contact author and guarantor of this review.

DECLARATIONS OF INTEREST

Mathilde Nijkeuter and Saskia Middeldorp were investigators in one included study (Nijkeuter 2013). These review authors were not involved in study selection nor in data extraction for this review.

TvM: none known.
LS: none known.
PdJ: none known.
ML: none known.
SM: Dr Middeldorp's institution has received funding for research grants from BMS/Pfizer, Daiichi Sankyo and Sanquin for investigator-initiated studies in the field of treatment of VTE and assessment of long-term outcomes of VTE; payment for lectures from GSK, Bayer, Boehringer Ingelheim and Sanofi; and payment for development of educational presentations from Bayer, GSK, Daiichi Sankyo and BMS/Pfizer. These companies had no influence on the content of the educational material.

SOURCES OF SUPPORT

Internal sources

- Academic Medical Center, Netherlands.

External sources

- Chief Scientist Office, Scottish Government Health Directorates, Scottish Government, UK.

  The Cochrane Vascular editorial base is supported by the Chief Scientist Office

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Additions to the protocol

1. We added a minimum sample size of fewer than 25 pregnant women per index test as an exclusion criterion. In this population, a PE is confirmed in 4% of women (Kline 2014). We therefore deemed cohorts of fewer than 25 as uninformative.
2. We regarded cases with an inconclusive index test who were subsequently treated on clinical grounds as having a positive reference standard. After an inconclusive index test, the chance of a PE is relatively high compared with after a negative test. Yet a missed PE is unlikely to be identified as such in the follow-up if a woman is receiving anticoagulant treatment. This would lead to an underestimation of false negatives. Therefore, the current adjudication was more conservative than the one provided in the original protocol.
3. We determined PE frequency. We considered this informative, but it was not specified in the protocol.

Alterations to the protocol

1. We removed the question in the QUADAS-2 checklist "If a threshold was used, was it prespecified?" because very few studies described the use of a threshold, and a prespecified threshold is not of the utmost importance when these imaging tests are performed.

Clarifications of the protocol

The original protocol did not provide guidance on all encountered judgement calls in data collection and analysis. We made the following clarifications.

1. CTPA, lung scintigraphy or MRA had to be performed as an index test, not as part of a reference test.
2. When data on both original assessments of the index test result were reported in the clinical setting, and when re-assessments for study purposes were reported, we preferred the former because this information matches the clinical question more accurately. We deemed re-assessments as having high risk of bias.
3. We set the maximum duration of the period between confirmed PE and the index test at three months. With later identified cases, we deemed the chance of newly occurring PE as substantial.
4. If patient recruitment was consecutive, but patients received one of multiple index tests on a non-systematic basis, we still regarded studies as presenting a consecutive patient series.
5. Lung scintigraphy as an index test comprised perfusion scintigraphy - planar or V/Q SPECT and with or without ventilation scanning.
6. When possible, we included only data on the first index test if the patient had received the same index test more than once. If followed by a different index test, we included this test result.

I N D E X   T E R M S

Medical Subject Headings (MeSH)

Angiography [standards] [statistics & numerical data]; Magnetic Resonance Angiography; Positron-Emission Tomography [*standards] [statistics & numerical data]; Pregnancy Complications, Hematologic [*diagnostic imaging]; Pulmonary Embolism [*diagnostic imaging]; Radionuclide Imaging [standards] [statistics & numerical data]; Sensitivity and Specificity; Tomography, X-Ray Computed [*standards] [statistics & numerical data]

MeSH check words

Female; Humans; Pregnancy