Cochrane Library

Ultrasound for fetal assessment in early pregnancy (Review)

Whitworth M, Bricker L, Mullan C

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Ultrasound for fetal assessment in early pregnancy

Melissa Whitworth[1], Leanne Bricker[2], Clare Mullan[3]

[1]St Mary's Hospital, Manchester, UK. [2]Corniche Hospital, Abu Dhabi, United Arab Emirates. [3]Department of Obstetrics & Gynaecology, St Mary's Hospital, Manchester, UK

Contact address: Melissa Whitworth, St Mary's Hospital, Central Manchester and Manchester Children's University Hospitals NHS Trust, Hathersage Road, Manchester, M13 0JH, UK. mkw@doctors.org.uk.

A B S T R A C T

Background

Diagnostic ultrasound is a sophisticated electronic technology, which utilises pulses of high-frequency sound to produce an image. Diagnostic ultrasound examination may be employed in a variety of specific circumstances during pregnancy such as after clinical complications, or where there are concerns about fetal growth. Because adverse outcomes may also occur in pregnancies without clear risk factors, assumptions have been made that routine ultrasound in all pregnancies will prove beneficial by enabling earlier detection and improved management of pregnancy complications. Routine screening may be planned for early pregnancy, late gestation, or both. The focus of this review is routine early pregnancy ultrasound.

Objectives

To assess whether routine early pregnancy ultrasound for fetal assessment (i.e. its use as a screening technique) influences the diagnosis of fetal malformations, multiple pregnancies, the rate of clinical interventions, and the incidence of adverse fetal outcome when compared with the selective use of early pregnancy ultrasound (for specific indications).

Search methods

We searched the Cochrane Pregnancy and Childbirth Group's Trials Register (30 March 2015) and reference lists of retrieved studies.

Selection criteria

Published, unpublished, and ongoing randomised controlled trials that compared outcomes in women who experienced routine versus selective early pregnancy ultrasound (i.e. less than 24 weeks' gestation). We have included quasi-randomised trials.

Data collection and analysis

Two review authors independently assessed trials for inclusion and risk of bias, extracted data and checked them for accuracy. We used the Review Manager software to enter and analyse data.

Main results

Routine/revealed ultrasound versus selective ultrasound/concealed: 11 trials including 37,505 women. Ultrasound for fetal assessment in early pregnancy reduces the failure to detect multiple pregnancy by 24 weeks' gestation (risk ratio (RR) 0.07, 95% confidence interval (CI) 0.03 to 0.17; participants = 295; studies = 7), *moderate quality of evidence*). Routine scans improve the detection of major fetal abnormality before 24 weeks' gestation (RR 3.46, 95% CI 1.67 to 7.14; participants = 387; studies = 2,*moderate quality of evidence*). Routine scan is associated with a reduction in inductions of labour for 'post term' pregnancy (RR 0.59, 95% CI 0.42 to 0.83; participants = 25,516; studies = 8), but the evidence related to this outcome is of low quality, because most of the pooled effect was provided by studies with design limitation with presence of heterogeneity (I^2 = 68%). Ultrasound for fetal assessment in early pregnancy does not impact on perinatal death (defined

as stillbirth after trial entry, or death of a liveborn infant up to 28 days of age) (RR 0.89, 95% CI 0.70 to 1.12; participants = 35,735; studies = 10, *low quality evidence*). Routine scans do not seem to be associated with reductions in adverse outcomes for babies or in health service use by mothers and babies. Long-term follow-up of children exposed to scan in utero does not indicate that scans have a detrimental effect on children's physical or cognitive development.

The review includes several large, well-designed trials but lack of blinding was a problem common to all studies and this may have an effect on some outcomes. The quality of evidence was assessed for all review primary outcomes and was judged as moderate or low. Downgrading of evidence was based on including studies with design limitations, imprecision of results and presence of heterogeneity.

Authors' conclusions

Early ultrasound improves the early detection of multiple pregnancies and improved gestational dating may result in fewer inductions for post maturity. Caution needs to be exercised in interpreting the results of aspects of this review in view of the fact that there is considerable variability in both the timing and the number of scans women received.

P L A I N   L A N G U A G E   S U M M A R Y

Routine compared with selective ultrasound in early pregnancy

Ultrasound is an electronic technology that uses the reflection of pulses of high-frequency sound to produce an image. Ultrasound may be used in a variety of circumstances during pregnancy. It has been assumed that the routine use of ultrasound in early pregnancy will result in the earlier detection of problems and improved management of pregnancy complications when compared with selective use for specific indications such as after clinical complications (e.g. bleeding in early pregnancy), or where there are concerns about fetal growth.

The focus of this review is routine early pregnancy ultrasound (before 24 weeks). We included 11 randomised controlled trials involving 37,505 women. Early ultrasound improved the early detection of multiple pregnancies and improved gestational dating, which may result in fewer inductions for post maturity. The detection of fetal malformation was addressed in detail in only two of the trials. There was no evidence of a significant difference between the screened and control groups for perinatal death. Results do not show that routine scans reduce adverse outcomes for babies or lead to less health service use by mothers and babies. Long-term follow-up of children exposed to scans before birth did not indicate that scans have a detrimental effect on children's physical or intellectual development. Studies were carried out over three decades and technical advances in equipment, more widespread use of ultrasonography, and increased training and expertise of operators may have resulted in more effective sonography.

Summary of findings for the main comparison.   Routine/revealed compared with selective/concealed ultrasound in early pregnancy for fetal assessment in early pregnancy

Routine/revealed compared with selective/concealed ultrasound in early pregnancy for fetal assessment in early pregnancy						

Patient or population: Pregnant women at less than 24 weeks' gestation undergoing fetal assessment by ultrasound
Settings: Norway, United Kingdom, United States and Sweden
Intervention: Routine/revealed
Comparison: Selective/concealed ultrasound in early pregnancy

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	selective/concealed ultrasound in early pregnancy	Routine/revealed				
Detection of fetal abnormality before 24 weeks' gestation	Study population		RR 3.46 (1.67 to 7.14)	387 (2 RCTs)	⊕⊕⊕⊝ MODERATE [1]	
	44 per 1000	153 per 1000 (74 to 316)				
	Moderate					
	25 per 1000	85 per 1000 (41 to 175)				
Detection of multiple pregnancy by 24 to 26 weeks' gestation (number NOT detected)	Study population		RR 0.07 (0.03 to 0.17)	295 (7 RCTs)	⊕⊕⊕⊝ MODERATE [1]	
	394 per 1000	28 per 1000 (12 to 67)				
	Moderate					
	250 per 1000	18 per 1000 (8 to 43)				
Induction of labour for 'post-term' pregnancy	Study population		RR 0.59 (0.42 to 0.83)	25516 (8 RCTs)	⊕⊕⊝⊝ LOW [1,2]	
	31 per 1000	18 per 1000 (13 to 25)				

	Moderate				
	38 per 1000	22 per 1000 (16 to 32)			
Perinatal death (all babies)	Study population		RR 0.89 (0.70 to 1.12)	35735 (10 RCTs)	⊕⊕⊝⊝ LOW [1,3]
	8 per 1000	7 per 1000 (6 to 9)			
	Moderate				
	10 per 1000	9 per 1000 (7 to 11)			

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **RR:** risk ratio

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1]Most of the pooled effect provided by studies with design limitations, including poor reporting of allocation concealment methods (-1).
[2]Presence of heterogeneity (I^2 = 68%) (-1).
[3]Wide CI crossing the line of no effect (-1).

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

Cochrane Database of Systematic Reviews

BACKGROUND

Diagnostic ultrasound examination may be employed in a variety of specific circumstances during pregnancy, such as after clinical complications, or where there are concerns about fetal growth. Because adverse outcomes may also occur in pregnancies without clear risk factors, assumptions have been made that routine ultrasound in all pregnancies will prove beneficial by enabling earlier detection and improved management of pregnancy complications. Routine screening may be planned for early pregnancy, late gestation, or both. The focus of this review is routine early pregnancy ultrasound.

Description of the intervention

Diagnostic ultrasound is a sophisticated electronic technology that utilises pulses of high-frequency sound to produce an image. A transducer, which is moved across the area to be examined, emits pulses of ultrasound that propagate through the tissues. Some pulses are reflected back to the transducer, which converts these returning echoes into electronic signals. The strength of the returning echo is determined by tissue interface characteristics. Returning signals are processed by a computer that displays each echo in both strength and position as an image on a screen. The quality of ultrasound imaging is dependent not only on the technical capabilities of the ultrasound equipment, but also on the experience and expertise of the operator; and standards are variable.

Diagnostic ultrasound examination may be employed in a variety of specific circumstances during pregnancy, such as: after clinical complications (e.g. bleeding in early pregnancy); where the fetus is perceived to be at particularly high risk of malformation; and where there are concerns regarding fetal growth. Because adverse outcomes may also occur in pregnancies without clear risk factors, assumptions have been made that the routine use of ultrasound in all pregnancies will prove beneficial. The rationale for such screening would be the detection of clinical conditions that place the fetus or mother at high risk, which would not necessarily have been detected by other means such as clinical examination, and for which subsequent management would improve perinatal outcome. Routine screening examinations may be planned for early pregnancy, late gestation, or both. The focus of this review is routine early pregnancy ultrasound; late pregnancy screening has been addressed in another Cochrane review (Bricker 2008).

The use of ultrasound to identify women at risk of preterm delivery by assessment of the cervix may be a component of screening before 24 weeks; this is outside the remit of this review and is considered elsewhere (Berghella 2013; Crane 2008).

How the intervention might work

Early Pregnancy complications and serum screening

An ultrasound at the time of antenatal booking may enable non-viable pregnancies to be detected earlier than is possible using clinical presentation. This has implications for clinical management of these pregnancies. In addition, earlier detection of ectopic pregnancy may be possible allowing for medical rather than surgical management, or 'minimally invasive' rather than open surgery. Between 11% and 42% of gestational age estimations taken from the menstrual history are reported as inaccurate (Barrett 1991; Geirsson 1991; Peek 1994). A reliable estimate of gestational age is required for maternal serum screening for fetal abnormality to be accurately timed (Owen 1997). Accurate knowledge of gestational age may increase the efficiency of maternal serum screening and some late pregnancy fetal assessment tests.

Multiple Pregnancy

Multiple pregnancies are associated with increased perinatal morbidity and mortality compared to singleton pregnancies (Dodd 2005). Determination of chorionicity plays an important role in risk stratification when managing twin pregnancies. Routine early pregnancy scanning in this group may impact on accuracy of assignment of chorionicity in multiple pregnancies, as some studies have shown that this can be done more accurately at earlier gestations (Lee 2006). It is also possible that earlier diagnosis of multiple pregnancy will occur with routine early pregnancy scanning, thus preventing inappropriate maternal serum screening (Persson 1983; Saari-Kemppainen 1990).

Detection of structural fetal abnormalities

In a systematic review, based on 11 studies (one randomised controlled trial, six retrospective cohorts and four prospective cohorts), undertaken to examine the use of routine second trimester ultrasound to detect fetal anomalies, the overall prevalence of fetal anomaly was 2.09%, ranging from 0.76% to 2.45% in individual studies and including major and minor anomalies (Bricker 2000). Using late pregnancy ultrasound scanning overall, detection of fetal anomaly was 44.7%, with a range of 15.0% to 85.3% (Bricker 2000). Optimum timing of such ultrasound scans may be aided by accurate estimation of dates using routine early pregnancy scanning.

Timing of delivery

A recent Cochrane review concluded that compared with a policy of expectant management, a policy of labour induction is associated with fewer perinatal deaths (Gulmezoglu 2012). It is possible that routine early pregnancy scanning will improve the accuracy of pregnancy dating and thereby affect the number of pregnancies undergoing induction for post-maturity. Whilst there is evidence to suggest that ultrasound is very attractive to women and families, studies have also shown that women often lack information about the purposes for which an ultrasound scan is being done and the technical limitations of the procedure (Bricker 2000). It is therefore essential that women's satisfaction is considered.

Why it is important to do this review

The use of routine pregnancy ultrasound needs to be considered in the context of potential hazards. In theory, some ultrasonic energy propagated through tissue is converted to heat, and in laboratory experiments, biological effects of ultrasound have been observed. However, these effects have been produced using continuous wave ultrasound with long 'dwell' time (time insonating one area) and high-power output. In the clinical setting, diagnostic ultrasound uses pulsed waves (short pulses of sound propagation), and most modern machines are designed so that safe power output limits cannot be exceeded. Operators are always advised to apply the ALARA (as low as reasonably attainable) principle to the ultrasound power output used (EFSUMB 1995), and to ensure time taken for an examination, including the 'dwell' time over a specific target, is kept to a minimum. One of the aims of this review is to assess available

data and determine whether clear epidemiological evidence exists from clinical trials that ultrasound examination during pregnancy is harmful.

OBJECTIVES

To assess whether routine early pregnancy ultrasound for fetal assessment (i.e. its use as a screening technique) influences the diagnosis of fetal malformations, multiple pregnancies, the rate of clinical interventions, and the incidence of adverse fetal outcome when compared with the selective use of early pregnancy ultrasound (for specific indications).

METHODS

Criteria for considering studies for this review

Types of studies

All published, unpublished, and ongoing randomised controlled trials with reported data that compared outcomes in women who experienced routine early pregnancy ultrasound with outcomes in women who experienced the selective use of early pregnancy ultrasound. We included quasi-randomised controlled trials. Cluster-randomised trials were also eligible for inclusion.

We planned to include trials reported as abstracts provided that they contained sufficient information for us to assess eligibility and risk of bias, and that results were described in sufficient detail. Where insufficient information was provided in abstracts, we have included studies in Characteristics of studies awaiting classification until publication of the full study report or until we can obtain further information from authors.

Types of participants

Women with early pregnancies, i.e. less than 24 weeks' gestation.

Types of interventions

Routine ultrasound examination compared with selective ultrasound examination.

Types of outcome measures

Primary outcomes

1. Detection of major fetal abnormality (as defined by the trial authors) prior to 24 weeks' gestation.
2. Detection of multiple pregnancy by 24 weeks' gestation.
3. Induction of labour for 'post-term' pregnancy.
4. Perinatal death (defined as stillbirth after trial entry, or death of a liveborn infant up to 28 days of age).

Secondary outcomes

Detection

1. Non-viable pregnancy prior to clinical presentation.
2. Ectopic pregnancy prior to clinical presentation.
3. Chorionicity of multiple pregnancy (in first trimester or in second trimester).
4. Multiple pregnancy prior to labour.
5. Soft markers before 24 weeks' gestation (i.e. structural features in the fetus that are of little or no functional significance (e.g. choroid plexus cyst, echogenic bowel), but which can

be associated with increased risk of chromosomal disorder, e.g.Trisomy 21).
6. Major anomaly before birth.

Complications for infants and children

1. Birthweight.
2. .Gestation at delivery.
3. Low birthweight (defined as less than 2500 g at term in singletons).
4. Very low birthweight (defined as less than 1500 g at term in singletons).
5. Apgar score less than or equal to seven at five minutes.
6. Admission to neonatal intensive care unit (NICU).
7. Respiratory distress syndrome.
8. Death or major neurodevelopmental handicap at childhood follow-up.
9. Poor oral reading at school.
10. Poor reading comprehension at school.
11. Poor spelling at school.
12. Poor arithmetic at school.
13. Poor overall school performance.
14. Dyslexia.
15. (Reduced hearing in childhood.
16. Reduced vision in childhood.
17. Use of spectacles.
18. Non right-handedness.
19. Ambidexterity.
20. Disability at childhood follow-up.

Maternal outcomes

1. Appropriately timed serum screening tests.
2. Laparoscopic management of ectopic pregnancy.
3. Surgical management of abortion.
4. Appropriately timed anomaly scan (18 to 22 weeks).
5. Termination of pregnancy for fetal abnormality.
6. Antenatal hospital admission.
7. Induction of labour for any reason.
8. Caesarean section.

Measures of satisfaction

1. Woman not satisfied.
2. Women's preferences for care.

Costs

1. Costs associated with routine early pregnancy ultrasound versus selective early pregnancy ultrasound.
2. Number of antenatal visits.
3. Length of stay in NICU.
4. Infant length of hospital stay.

Search methods for identification of studies

The following methods section of this review is based on a standard template used by the Cochrane Pregnancy and Childbirth Group.

Electronic searches

We searched the Cochrane Pregnancy and Childbirth Group's Trials Register by contacting the Trials Search Co-ordinator (30 March 2015).

The Cochrane Pregnancy and Childbirth Group's Trials Register is maintained by the Trials Search Co-ordinator and contains trials identified from:

1. monthly searches of the Cochrane Central Register of Controlled Trials (CENTRAL);
2. weekly searches of MEDLINE (Ovid);
3. weekly searches of Embase (Ovid);
4. monthly searches of CINAHL (EBSCO);
5. handsearches of 30 journals and the proceedings of major conferences;
6. weekly current awareness alerts for a further 44 journals plus monthly BioMed Central email alerts.

Details of the search strategies for CENTRAL, MEDLINE, Embase and CINAHL, the list of handsearched journals and conference proceedings, and the list of journals reviewed via the current awareness service can be found in the 'Specialized Register' section within the editorial information about the Cochrane Pregnancy and Childbirth Group.

Trials identified through the searching activities described above are each assigned to a review topic (or topics). The Trials Search Co-ordinator searches the register for each review using the topic list rather than keywords.

Searching other resources

We examined cited references, abstracts, letters to the editor, and editorials for additional studies. Where necessary, we contacted the primary investigator directly to obtain further data.

We did not apply any language or date restrictions.

Data collection and analysis

For methods used in the previous version of this review, *see* Whitworth 2010.

For this update, the following methods were used for assessing the six new reports that were identified as a result of the updated search and the four reports in the 'Studies awaiting classification' of the previous version of the review (Whitworth 2010).

The following methods section of this review is based on a standard template used by the Cochrane Pregnancy and Childbirth Group.

Selection of studies

Two review authors independently assessed for inclusion all the potential studies identified as a result of the search strategy. We resolved any disagreement through discussion or, if required, we consulted the third review author.

Data extraction and management

We designed a form to extract data. For eligible studies, two review authors extracted the data using the agreed form. We resolved discrepancies through discussion or, if required, we consulted the

third review author. We entered data into Review Manager software (RevMan 2014) and checked for accuracy.

When information regarding any of the above was unclear, we contacted authors of the original reports to provide further details.

Assessment of risk of bias in included studies

Two review authors independently assessed risk of bias for each study using the criteria outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). Any disagreement was resolved by discussion or by involving a third assessor.

(1) Random sequence generation (checking for possible selection bias)

We described for each included study the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.

We assessed the method as:

- low risk of bias (any truly random process, e.g. random number table; computer random number generator);
- high risk of bias (any non-random process, e.g. odd or even date of birth; hospital or clinic record number);
- unclear risk of bias.

(2) Allocation concealment (checking for possible selection bias)

We described for each included study the method used to conceal allocation to interventions prior to assignment and assessed whether intervention allocation could have been foreseen in advance of, or during recruitment, or changed after assignment.

We assessed the methods as:

- low risk of bias (e.g. telephone or central randomisation; consecutively numbered sealed opaque envelopes);
- high risk of bias (open random allocation; unsealed or non-opaque envelopes, alternation; date of birth);
- unclear risk of bias.

(3.1) Blinding of participants and personnel (checking for possible performance bias)

We described for each included study the methods used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. We considered that studies were at low risk of bias if they were blinded, or if we judged that the lack of blinding unlikely to affect results. We assessed blinding separately for different outcomes or classes of outcomes.

We assessed the methods as:

- low, high or unclear risk of bias for participants;
- low, high or unclear risk of bias for personnel.

(3.2) Blinding of outcome assessment (checking for possible detection bias)

We described for each included study the methods used, if any, to blind outcome assessors from knowledge of which intervention a participant received. We assessed blinding separately for different outcomes or classes of outcomes.

We assessed methods used to blind outcome assessment as:

- low, high or unclear risk of bias.

(4) Incomplete outcome data (checking for possible attrition bias due to the amount, nature and handling of incomplete outcome data)

We described for each included study, and for each outcome or class of outcomes, the completeness of data including attrition and exclusions from the analysis. We stated whether attrition and exclusions were reported and the numbers included in the analysis at each stage (compared with the total randomised participants), reasons for attrition or exclusion where reported, and whether missing data were balanced across groups or were related to outcomes. Where sufficient information was reported, or could be supplied by the trial authors, we planned to re-include missing data in the analyses which we undertook.

We assessed methods as:

- low risk of bias (e.g. no missing outcome data; missing outcome data balanced across groups);
- high risk of bias (e.g. numbers or reasons for missing data imbalanced across groups; 'as treated' analysis done with substantial departure of intervention received from that assigned at randomisation);
- unclear risk of bias.

(5) Selective reporting (checking for reporting bias)

We described for each included study how we investigated the possibility of selective outcome reporting bias and what we found.

We assessed the methods as:

- low risk of bias (where it is clear that all of the study's pre-specified outcomes and all expected outcomes of interest to the review have been reported);
- high risk of bias (where not all the study's pre-specified outcomes have been reported; one or more reported primary outcomes were not pre-specified; outcomes of interest are reported incompletely and so cannot be used; study fails to include results of a key outcome that would have been expected to have been reported);
- unclear risk of bias.

(6) Other bias (checking for bias due to problems not covered by (1) to (5) above)

We described for each included study any important concerns we had about other possible sources of bias.

(7) Overall risk of bias

We made explicit judgements about whether studies were at high risk of bias, according to the criteria given in the *Handbook* (Higgins 2011). With reference to (1) to (6) above, we planned to assess the likely magnitude and direction of the bias and whether we considered it is likely to impact on the findings. We explored the impact of the level of bias through undertaking sensitivity analyses - *see* Sensitivity analysis.

Assessment of the quality of the body of evidence

For this update, we assessed the quality of the evidence using the GRADE approach (Schunemann 2009). We assessed the quality of the body of evidence relating to the following outcomes for the main comparison of routine early ultrasound scan versus no early ultrasound.

1. Detection of major fetal abnormality (as defined by the trial authors) prior to 24 weeks' gestation.
2. Detection of multiple pregnancy by 24 weeks' gestation.
3. Induction of labour for 'post-term' pregnancy.
4. Perinatal death (defined as stillbirth after trial entry, or death of a liveborn infant up to 28 days of age).

We used GRADEprofiler (GRADEpro 2014) to import data from Review Manager 5.3 (RevMan 2014) in order to create a 'Summary of findings' table. We produced a summary of the intervention effect and a measure of quality for each of the above outcomes using the GRADE approach. The GRADE approach uses five considerations (study limitations, consistency of effect, imprecision, indirectness and publication bias) to assess the quality of the body of evidence for each outcome. The evidence can be downgraded from 'high quality' by one level for serious (or by two levels for very serious) limitations, depending on assessments for risk of bias, indirectness of evidence, serious inconsistency, imprecision of effect estimates or potential publication bias.

Measures of treatment effect

Dichotomous data

For dichotomous data, we presented results as summary risk ratio with 95% confidence intervals.

Continuous data

We used the mean difference if outcomes were measured in the same way between trials. In future updates, if the data dictate, we will use the standardised mean difference to combine trials that measure the same outcome, but use different methods.

Unit of analysis issues

Cluster-randomised trials

We did not include cluster-randomised trials in this update. If in future cluster-randomised trials are identified, we will include them in the analyses along with individually-randomised trials. We will adjust their sample sizes using the methods described in the Handbook *[Section 16.3.4 or 16.3.6]* using an estimate of the intracluster correlation co-efficient (ICC) derived from the trial (if possible), from a similar trial or from a study of a similar population. If we use ICCs from other sources, we will report this and conduct sensitivity analyses to investigate the effect of variation in the ICC. If we identify both cluster-randomised trials and individually-randomised trials, we plan to synthesise the relevant information. We will consider it reasonable to combine the results from both if there is little heterogeneity between the study designs and the interaction between the effect of intervention and the choice of randomisation unit is considered to be unlikely.

We will also acknowledge heterogeneity in the randomisation unit and perform a sensitivity analysis to investigate the effects of the randomisation unit.

Cross-over trials

Cross-over trials were not eligible for inclusion in this review.

Other unit of analysis issue

We have included multiple pregnancies in this review but we have not done any adjustment. Multiple pregnancy data were analysed along side singleton births.

Dealing with missing data

For included studies, we noted levels of attrition. In future updates, if more eligible studies are included, we will explore the impact of including studies with high levels of missing data in the overall assessment of treatment effect by using sensitivity analysis.

For all outcomes, we carried out analyses, as far as possible, on an intention-to-treat basis, i.e. we attempted to include all participants randomised to each group in the analyses. The denominator for each outcome in each trial was the number randomised minus any participants whose outcomes were known to be missing.

Assessment of heterogeneity

We assessed statistical heterogeneity in each meta-analysis using the Tau^2, I^2 and Chi^2 statistics. We regarded heterogeneity as substantial if an I^2 was greater than 30% and either the Tau^2 was greater than zero, or there was a low P value (less than 0.10) in the Chi^2 test for heterogeneity. If we identified substantial heterogeneity (above 30%), we planned to explore it by pre-specified subgroup analysis.

Assessment of reporting biases

We planned to investigate reporting biases (such as publication bias) using funnel plots, if there were 10 or more studies in the analysis. One analysis did include 10 studies (perinatal death). For this outcome, we assessed funnel plot asymmetry visually. If asymmetry was suggested by a visual assessment, we planned to perform exploratory analyses to investigate it.

Data synthesis

We carried out statistical analysis using the Review Manager software (RevMan 2014). We used fixed-effect meta-analysis for combining data where it was reasonable to assume that studies were estimating the same underlying treatment effect: i.e. where trials were examining the same intervention, and the trials' populations and methods were judged sufficiently similar.

If there was clinical heterogeneity sufficient to expect that the underlying treatment effects differed between trials, or if substantial statistical heterogeneity was detected, we used random-effects meta-analysis to produce an overall summary, if an average treatment effect across trials was considered clinically meaningful. The random-effects summary was treated as the average of the range of possible treatment effects and we discussed the clinical implications of treatment effects differing between trials. If the average treatment effect was not clinically meaningful, we did not combine trials. Where we used random-effects analyses, the results were presented as the average treatment effect with 95% confidence intervals, and the estimates of Tau^2 and I^2.

Subgroup analysis and investigation of heterogeneity

If we identified substantial heterogeneity, we investigated it using subgroup analyses and sensitivity analyses. We considered whether an overall summary was meaningful, and if it was, we used random-effects analysis to produce it.

Planned subgroup analyses included:

1. parity (nulliparous versus multiparous women);
2. selective performance of ultrasound versus selective reporting of ultrasound findings;
3. first scan occurring in first trimester (up to 14 weeks' gestation) versus second trimester (14 to 24 weeks' gestation).

We carried out subgroup analyses for the review's primary outcomes (induction for post-term pregnancy, perinatal death, and detection of multiple pregnancy and abnormality before 24 weeks' gestation). We assessed differences between subgroups by visual inspection of the forest plots and the subgroups' confidence intervals; non-overlapping confidence intervals indicating a statistically significant difference in treatment effect between the subgroups. If we suspected any differences between subgroups we planned to seek statistical advice.

We assessed subgroup differences by interaction tests available within RevMan (RevMan 2014). We reported the results of subgroup analyses quoting the Chi^2 statistic and P value, and the interaction test I^2 value.

Sensitivity analysis

We planned to carry out sensitivity analyses to explore the effect of trial quality assessed by concealment of allocation, high attrition rates, or both, with poor quality studies being excluded from the analyses in order to assess whether this makes any difference to the overall result.

R E S U L T S

Description of studies

Results of the search

For the previous version of this review (Whitworth 2010), we identified 63 papers reporting findings from 24 studies examining ultrasound for fetal assessment in early pregnancy (most studies resulted in several publications or reports). We included 11 trials involving 37,505 women. One study had no published data (Newcastle 2004); we contacted the authors of this study and added it to Studies awaiting classification. In addition, one of the included studies (Norway 1993) reported long-term, childhood outcome data from two of the included trials (Alesund 1984; Trondheim 1984). We excluded 11 studies.

For this current version of the review, an updated search in March 2015 identified six new reports. One new trial (McClure 2014), has been added to ongoing studies. We identified one additional report for Sweden 1988 and one additional report for Johannesburg 2007. Three reports of new studies have been excluded (Georgsson Ohman 2010; Schifano 2010; Votino 2012). Newcastle 2004 still has no published data and therefore is still awaiting assessment.

Included studies

The studies were carried out in a number of countries: Australia (Adelaide 1999); the USA (RADIUS 1993 (a multicentre study); Missouri 1990); South Africa (Johannesburg 2007; Tygerberg 1996); Sweden (Sweden 1988); Norway (Alesund 1984; Trondheim 1984); the UK (London 1982; Oxford 2006) and Finland (Helsinki 1990). The earliest trials began recruitment in the late 1970s (Alesund 1984; Trondheim 1984).

All of the trials included an intervention involving an ultrasound examination before the 24th week of pregnancy. The dates of the scans, and the number of scans women received varied in different trials. In the London 1982 trial, all women (in both the intervention and control groups) were offered a scan, but while in the intervention group results were revealed in the women's case notes, in the control group results were concealed unless they were specifically requested by clinical staff. In all other included studies, women in the intervention group were offered a "routine" scan whilst those in the control groups received a scan at the discretion of clinical staff ("selective scans"). Ultrasound scans in the intervention group may have been the only 'routine' scan offered, or may have been an additional scan, with women in both intervention and control groups having scans scheduled at a later stage of pregnancy.

The gestational age at which index scans were performed, and the purpose of scans, varied in different trials. In the Adelaide 1999 study, scans in the intervention group were carried out at between 11 and 14 weeks. The purposes of the scan were to ascertain gestational age (with the aim of improving the timing of other screening tests), identify multiple pregnancies, and to carry out a limited examination of fetal morphology. Women in both treatment and control arms of this study had a routine morphology scan at 18 to 20 weeks' gestation.

In the two Norwegian studies, women in the intervention group were offered two scans, the first at 18 to 20 weeks and then a late ultrasound scan at 32 weeks (Alesund 1984; Trondheim 1984), with follow-up data for both studies reported in the Norway 1993 papers. The purposes of the early scan were to measure biparietal diameter (BPD); to estimate the expected date of delivery (EDD); to identify multiple pregnancies; to note the location of the placenta; and to carry out a general examination of the fetus.

The ultrasound in the intervention group of the Helsinki 1990 trial was carried out between 16 and 20 weeks; the aims were similar to those in the Norwegian trials, with the amount of amniotic fluid also being recorded.

The main aim in the London 1982 study was BPD measurement and assessment of the EDD; scans were performed at approximately 16 weeks. In the Missouri 1990 trial, scans generally took place between 10 to 12 weeks (up to 18 weeks) and were carried out to estimate gestational age, identify multiple pregnancies, assess fetal viability, and to identify uterine abnormalities. The Sweden 1988 study had similar aims; women attending 19 antenatal clinics in Stockholm were invited for a scan at 15 weeks (range 13 to 19 weeks) (intervention) or received selective scans after 19 weeks (control).

In the Oxford 2006 study, women in both the intervention and control groups had routine scans at 18 to 20 weeks; in addition, women in the intervention group were offered an early scan (at between eight and 12 weeks) to estimate gestational age.

In a large multicentre trial in the USA (RADIUS 1993), women in the intervention group were offered scans at 18 to 20 weeks and at 31 to 33 weeks versus selective scans. The purpose of the earlier scan was to identify the location of the placenta, the volume of amniotic fluid, uterine abnormalities, multiple pregnancies, BPD and other measures of fetal size, and a detailed assessment of fetal anatomy.

In the South African study (Tygerberg 1996), a single scan was offered to women in the intervention group; women received a scan which aimed to ascertain gestational age and to identify major fetal anomalies. Finally, in the Johannesburg 2007 trial, scans in the intervention group were carried out between 18 and 23 weeks. The purpose of the scan was determination of single or multiple pregnancy; placental site identification; estimation of gestation age based on a combination of biparietal diameter, head circumference and femur length measurements, and to search for fetal abnormalities.

Further details of settings, participants and interventions are set out in the Characteristics of included studies tables.

Excluded studies

Altogether, 14 studies have been excluded (Belanger 1996; Bennett 2004; Duff 1993; Georgsson Ohman 2010; Hong Kong; Larsen 1992; Leung 2006; Owen 1994; Rustico 2005; Saltvedt 2006; Schifano 2010; Schwarzler 1999; Votino 2012; Wald 1988).

In four of the excluded studies, all participants (in both the intervention and control groups) received early scans. In the studies by Saltvedt 2006 and Schwarzler 1999, the timing of scans was examined (i.e. earlier versus later scans); in the study by Duff 1993, women in the intervention group had an additional scan in the third trimester; and Owen 1994 looked at women with a high risk of fetal anomaly, with women in the intervention group receiving more frequent scans.

In Bennett 2004, the focus was specifically on the timing of the assessment of gestational age with scans in the first and second trimesters. In Larsen 1992, the participants were high-risk women, with those in the intervention group receiving an additional scan at 28 weeks. Two trials compared two versus three or four dimensional ultrasound scans (Leung 2006; Rustico 2005). In the Hong Kong study, women in both arms of the trial had two routine scans; in the intervention group the earlier scan was more detailed than in the control group. The study by Belanger 1996 did not include results relevant to the outcomes of the review. We contacted the authors of Wald 1988, but results were not available.

Risk of bias in included studies

See Figure 1 and Figure 2 for a summary of 'Risk of bias' assessments.

Figure 1.   'Risk of bias' summary: review authors' judgements about each risk of bias item for each included study.

Figure 2. **'Risk of bias' graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.**

Allocation

In the majority of the included studies, no information was provided on the methods used to generate the sequence for randomisation. In the Adelaide 1999 study, a table of random numbers was used to generate the allocation order, and in the RADIUS 1993 study, the sequence was determined by a computerised random number generator. In two studies there was balanced block randomisation (Missouri 1990 (block size four) and Oxford 2006 (block size six)).

We assessed two studies as having inadequate methods to conceal group allocation; in the London 1982 study there was quasi-random allocation to groups using the women's hospital number to determine allocation, while in the Johannesburg 2007 trial, women were allocated to intervention and control groups according to day of the week. In the RADIUS 1993 study, the methods used for allocating women to randomisation group were not clear. In all the remaining studies, the study allocation was concealed in sealed envelopes; in the Adelaide 1999, Missouri 1990 and Oxford 2006 studies envelopes were described as numbered, opaque, and sealed; in the Tygerberg 1996 and Sweden 1988 studies envelopes were sealed and opaque; while the Alesund 1984, Helsinki 1990, and Trondheim 1984 reports refer to the "sealed envelope method" of randomisation.

Blinding

Blinding women and clinical staff to group allocation was generally not feasible as women in the two treatment arms received different care, and the results of scans were recorded in women's case notes. In the Trondheim 1984 study outcome assessment was described as partially blinded. In the London 1982 study the results of the scan for the control group were not stored in case notes and therefore not available to outcome assessors (although outcome assessors would be aware of group allocation by the presence or absence of the scan report).

The lack of blinding in these studies is unlikely to affect some review outcomes (such as perinatal mortality), but outcomes relying on

clinical judgement (e.g. the decision to induce labour for post-term pregnancy) may possibly be influenced by knowledge of group allocation and this is a potential source of bias; such outcomes should be interpreted with caution.

Incomplete outcome data

The loss of women to follow-up and levels of missing data for pregnancy outcomes were generally low in these studies (less than 5%). In Missouri 1990, 9% of the sample were lost to follow-up after randomisation, and in the Oxford 2006 study 15% were lost, but an intention-to-treat analysis, including all women randomised, was carried out for the main study outcomes. There was relatively high attrition in Johannesburg 2007; 15.7% of women randomised were lost to follow-up and there were further missing data for some outcomes; reasons for attrition were not described, and it was not clear how many women from each group were lost.

We have attempted to use consistent denominators in the analyses within this review. For pregnancy and early postnatal outcomes, we tried to include all women randomised less: those women that were not actually pregnant; those who had pre-screening pregnancy termination; and those who experienced an early spontaneous miscarriage. We included women lost to follow-up for other reasons (for example, did not attend for screening, withdrew from the study, missing data) in the denominators. In some cases, it was difficult to determine the denominators, as detailed information on attrition at different stages was not reported (in the Johannesburg 2007 study it was not clear how many women were randomised to each study group, and so we had to use the group denominators for those women available to follow-up). For outcomes for "all babies", we have used the total number of babies (including babies from multiple pregnancies); some outcomes (e.g. birthweight) are specified for singletons only. Again, it was not always easy to ascertain the appropriate denominators for babies.

For long-term follow-up, where there was greater attrition (for example, there was complete data for approximately half of the original sample for some outcomes at childhood follow-up in the

Sweden 1988 study), we have used the denominators reported by the authors in study publications.

For some competing/overlapping outcomes (e.g. perinatal deaths, miscarriages and pregnancy termination for fetal abnormality) we have reported figures provided in the trial reports, but we advise caution in interpreting such data. We will return to this issue in the Discussion.

Selective reporting

We found no evidence of selective reporting. However, protocols were not available for us to assess this, so we do not know the original planned outcomes.

Other potential sources of bias

Some of the trials had other potential sources of bias: the Oxford 2006 study was stopped part way through and results are difficult to interpret, and in Alesund 1984, there was some baseline imbalance between groups in smoking rates. While not a source of bias as such, large numbers of women were not eligible for inclusion in the RADIUS 1993 trial and this may affect the generalisability of results.

Effects of interventions

See: **Summary of findings for the main comparison** Routine/ revealed compared with selective/concealed ultrasound in early pregnancy for fetal assessment in early pregnancy

Routine/revealed ultrasound versus selective ultrasound/ concealed results: 11 trials including 37,505 women

Primary outcomes

1. Detection of major fetal abnormality (as defined by the trial authors) prior to 24 weeks' gestation

The detection of fetal abnormalities before 24 weeks in the screened and unscreened groups was reported in two studies; these studies (17,158 pregnancies) recorded a total of 387 fetal abnormalities with most being undetected at 24 weeks (346, 89% not detected by 24 weeks). It was more likely for the screened group to have abnormalities detected by 24 weeks compared with controls (unweighted percentages 16% versus 4%; risk ratio (RR) 3.46, 95% confidence interval (CI) 1.67 to 7.14; participants = 387; studies = 2) (Analysis 1.1).

2. Detection of multiple pregnancy by 24 weeks' gestation

Failure to detect multiple pregnancies by 24 weeks was reported in seven studies. It was more likely that multiple pregnancies would not be detected by 24 weeks in the unscreened groups; only two of 153 multiple pregnancies were undetected at 24 weeks in the screened groups, compared with 56 of 142 in the control groups (RR 0.07, 95% CI 0.03 to 0.17; participants = 295; studies = 7) (Analysis 1.2).

3. Induction of labour for 'post-term' pregnancy

Eight studies reported rates of induction of labour for 'post-dates' pregnancy (which accounted for approximately 13% of total inductions). Compared with controls, women offered early routine ultrasound were less likely to be induced for post-maturity. For this outcome there was a high level of heterogeneity between studies. A visual examination of the forest plot reveals that the general direction of findings is the same among studies; however, the size

of the treatment effect and the rates of induction in control groups vary. We used a random-effects model in the meta-analysis and the average treatment effect favoured the screened group (average RR 0.59, 95% CI 0.42 to 0.83; participants = 25,516; studies = 8; I² = 68%) (Analysis 1.3).

4. Perinatal death (defined as stillbirth after trial entry, or death of a liveborn infant up to 28 days of age)

There was no evidence of a significant difference between the screened and control groups for perinatal mortality (unweighted percentages 0.73 versus 0.82%; RR 0.89, 95% CI 0.70 to 1.12; participants = 35,735; studies = 10) (Analysis 1.4). When lethal malformations were excluded, rates of perinatal death in the screened and unscreened groups were very similar (0.53 versus 0.56%), (RR 0.96, 95% CI 0.72 to 1.27; participants = 34,331; studies = 8) (Analysis 1.5).

Secondary outcomes

Detection of abnormalities and multiple pregnancies prior to labour

All multiple pregnancies (140) were detected before labour in the intervention groups, whereas 12 of the 133 in the control groups remained undetected at the onset of labour (RR 0.12, 95% CI 0.03 to 0.54; participants = 273; studies = 5) (Analysis 1.6). Screened groups were also more likely to have major fetal anomalies detected before birth (RR 3.19, 95% CI 1.99 to 5.11; participants = 387; studies = 2) (Analysis 1.7).

Complications for infants and children

There was no evidence of significant differences between groups in terms of the number of low birthweight babies (less than 2500 g) (average RR 1.04, 95% CI 0.82 to 1.33; participants = 19,337; studies = 8; I² = 67%) (Analysis 1.8), or very low birthweight babies (less than 1500 g); (average RR 1.26, 95% CI 0.27 to 5.82; participants = 1584; studies = 2; I² = 52%) (Analysis 1.9). For these outcomes, some studies reported results for singletons only, and so in the analyses we have set out results for singletons and all babies separately. There was no evidence of statistically significant differences between groups in the number of babies that were small-for-gestational age (average RR 1.05, 95% CI 0.81 to 1.35; participants = 17105; studies = 3; I² = 40%) (Analysis 1.10), or in mean birthweight (mean difference (MD) 10.67, 95% CI -19.77 to 41.11; participants = 23213; studies = 5; I² = 59%) (Analysis 1.11). There were high levels of heterogeneity for outcomes relating to low birthweight, and these results should be interpreted with caution.

The number of babies with low Apgar scores (seven or less) at five minutes was similar in the two groups (average RR 0.76, 95% CI 0.33 to 1.72; participants = 3906; studies = 4; I² = 44%)) (Analysis 1.12), and there was no difference in rates of admission to neonatal intensive care (RR 0.95, 95% CI 0.88 to 1.02; participants = 19,088; studies = 8) (Analysis 1.13).

In three studies, babies were followed up into childhood (results for children up to nine years in the Alesund 1984 and Trondheim 1984 trials are reported together in Norway 1993, and the Sweden 1988 study includes data for children at eight to nine years and for teenagers aged 15 to 16 years). For children aged eight to nine years, there were no significant differences for any of the outcomes reported including school performance, hearing and vision, disabilities or dyslexia (which was measured in a subset of the main sample) (Analysis 1.14; Analysis 1.15; Analysis 1.16;

Analysis 1.17; Analysis 1.18; Analysis 1.19; Analysis 1.20; Analysis 1.21; Analysis 1.22; Analysis 1.23; Analysis 1.24; Analysis 1.25).There was concern raised regarding an excess of non-right handedness in the intervention group in the Norwegian study; however, the Swedish study did not confirm these findings, and results may have occurred by chance. Examination of the school records of teenagers (aged 15 to 16 years) revealed little difference in the performance of children whose mothers had been randomised to ultrasound or no ultrasound in the Swedish trial. Data were available for 94% of singletons in the original sample. Authors reported that there was no strong evidence of differences between groups for school performance (grades over all subjects) for girls or boys. For physical education, there was a small difference in scores for boys; those whose mothers had been randomised to the ultrasound group had slightly lower mean scores compared to those in the control group, but this finding was not statistically significant.

Maternal outcomes

It was more likely that women would undergo pregnancy termination for fetal abnormality in the screened groups, although overall the number of terminations was small (24 of 14,237 pregnancies in screened groups were terminated after detection of abnormality compared with 10 of 14,019 in controls) (RR 2.23, 95% CI 1.10 to 4.54; participants = 28,256; studies = 5) (Analysis 1.28).

There was no significant evidence that ultrasound was associated with reduced numbers of women undergoing delivery by caesarean section (RR 1.05, 95% CI 0.98 to 1.12; participants = 22,193; studies = 5) (Analysis 1.32). Overall, on average, there were slightly fewer inductions of labour (for any reason, including post-maturity) in women in the screened groups; in view of heterogeneity we used a random-effects model for this outcome (RR 0.78, 95% CI 0.63 to 0.97; participants = 24790; studies = 7; I² = 84%) (Analysis 1.31). The rate of induction in the screened group was 18.8% versus 19.8% in the control group (unweighted percentages).

The Adelaide 1999 trial examined whether an early scan would reduce the number of serum screening tests or fetal anomaly scans that needed to be repeated because they had been performed at the wrong gestational age. There was no significant evidence that the numbers of women having repeat testing was reduced in the intervention group (RR 0.89, 95% CI 0.45 to 1.76; participants = 602; studies = 1) (Analysis 1.26); (RR 0.77, 95% CI 0.55 to 1.08; participants = 602; studies = 1) (Analysis 1.27). There was also no significant evidence in the Helsinki 1990 and Johannesburg 2007 trials that the number of antenatal visits was reduced (MD 0.16, 95% CI -0.33 to 0.65; participants = 9502; studies = 2; I² = 91%) (Analysis 1.29), and pooled results from five trials showed no significant reduction in antenatal hospital admissions (RR 1.04, 95% CI 0.92 to 1.18; participants = 17,785; studies = 6; I² = 67%) (Analysis 1.30).

In the Adelaide 1999, study investigators examined whether having an early scan was reassuring or worrying to mothers. Fewer mothers in the screened group reported feeling worried about their pregnancies (RR 0.80, 95% CI 0.65 to 0.99; participants = 634; studies = 1) (Analysis 1.33).

Costs of care

The impact of screening on costs to women and health services was examined in two trials.

In the Helsinki 1990 study, the average time spent in the hospital was 61 minutes and women spent 74 minutes travelling to hospital; 81% of the sample were working, and half of the working women used work time to attend for initial screening (some women may have had to attend for further screening-induced appointments).

The cost to health services was not simple to calculate; the cost of the examinations was offset by fewer hospital visits and stays, and in this study there was a lower perinatal mortality rate and increased pregnancy termination in the screened group, leading the authors to conclude that ultrasound resulted in cost savings (Leivo 1996).

The issue of cost was also examined in a trial carried out in a low-resource setting where overall adverse fetal outcome was higher, but where fetal anomalies represented a smaller proportion of adverse outcomes compared with high-resource settings. An explicit aim of the Tygerberg 1996 study was to examine whether the costs of routine ultrasound for all women (rather than selective ultrasound) would be offset by a reduction in the use of other healthcare resources. In this study, routine ultrasound was perceived as being an expensive luxury: "In our less privileged community, however, the cost of investigation is in direct competition with resources for more urgent needs in healthcare and housing, sanitation, education and unemployment ... a routine obstetric ultrasonography policy is expensive and ... the more selective use is not accompanied by an increase in adverse perinatal outcome" (Tygerberg 1996 p.507).

Other outcomes

Included studies did not report data for a number of the secondary outcomes pre-specified in our review protocol including the detection of ectopic pregnancy or chorionicity of multiple pregnancy, laparoscopic management of ectopic pregnancy and surgical management of abortion.

Subgroup and sensitivity analysis

We planned subgroup analysis by parity, by the timing of the early ultrasound (before or after 14 weeks) and by whether the control group had scans (with results concealed) rather than selective scans. We examined subgroups for primary outcomes only.

Parity

Information on parity was not available for us to be able to carry out this analysis.

Timing of early scan

In three studies the early scan was planned for before 14 weeks' gestation. In the Adelaide 1999 study, scans were planned for 11 to 14 weeks, in Oxford 2006 10 to 12 weeks, and in Missouri 1990, while scans could be performed up to 18 weeks, most were carried out between 10 and 12 weeks. There were no clear differences between subgroups where scans were performed earlier rather than later for perinatal death (Analysis 1.36). The detection of multiple pregnancy by 24 to 26 weeks' gestation appears better in those scanned after 14 weeks' gestation (Analysis 1.34), although only one study (Missouri 1990) reported this outcome for women scanned prior to 14 weeks with small event numbers. For detection of multiple pregnancy prior to 24 to 26 weeks' gestation In the Adelaide 1999 and Oxford 2006 studies, the treatment effect appeared more conservative than in some of the other studies; this

may be because women in both arms of these trials had routine ultrasound scheduled at 18 to 20 weeks' gestation. Compared with controls, women offered early routine ultrasound after 14 weeks' gestation were less likely to be induced for post-maturity (Analysis 1.35), whilst scans prior to 14 weeks' gestation had no impact.

Concealed results and routine care

In one study, women in both groups were screened but results were revealed for the intervention group only (London 1982). This study examined two of the review's primary outcomes: detection of multiple pregnancy before 24 weeks and perinatal death. There was considerable overlap in the confidence intervals for these outcomes London 1982 and the other trials, suggesting no clear differences between subgroups. There was no evidence

of a difference between subgroups according to the subgroup interaction test (Analysis 1.37; Analysis 1.38).

Sensitivity analysis (allocation concealment assessed as inadequate)

Two studies used a quasi-randomised design (case note number London 1982; allocation by day of the week Johannesburg 2007). Removing these studies from the analysis did not affect overall results for the primary outcomes. None of the studies had very high levels of attrition (more than 20%) for primary outcomes.

Only one outcome (perinatal death) included data from 10 studies (Analysis 1.4); we produced a funnel plot to look for plot asymmetry which can suggest publication bias; there was no asymmetry apparent on visual inspection (Figure 3).

Figure 3. Funnel plot of comparison: 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, outcome: 1.4 Perinatal death (all babies).

DISCUSSION

Summary of main results

Ultrasound for fetal assessment in early pregnancy increases the chances of the detection of multiple pregnancy before 24 weeks' gestation, and there is some evidence that fetal abnormalities are detected earlier. Routine scanning is associated with a reduction in inductions of labour for 'post-term' pregnancy, and this contributes to a small reduction in the overall rates of induction (for any indication). Routine scans do not seem to be associated with reductions in adverse outcomes for babies or in health service

use by mothers and babies. At the same time, long-term follow-up of children exposed to a scan in utero does not indicate that scans have a detrimental effect on children's physical or cognitive development.

Considerable caution needs to be exercised in interpreting the results of aspects of this review in view of the fact that there is considerable variability in both the timing of the intervention and the number of scans that women received during pregnancy.

The assumed benefits of routine ultrasonography in early pregnancy have been: (1) better gestational age assessment;

(2) earlier detection of multiple pregnancies; and (3) detection of clinically unsuspected fetal malformation at a time when termination of pregnancy is possible.

These assumptions appear to have been justified by analysis of data from the studies included in this review. The reduced incidence of induction of labour in the routinely scanned groups presumably results from better gestational 'dating', and earlier detection of multiple pregnancy. However, the high levels of heterogeneity for the former outcome means caution should be applied. Whilst routine ultrasound assessment in early pregnancy has not been shown to improve fetal outcome, much larger numbers of participants would be required to demonstrate that better gestational 'dating' and earlier detection of multiple pregnancy result in improved outcomes for babies.

The detection of fetal malformation has been addressed in detail only in two of the trials. The Helsinki 1990 trial showed improved detection with a resultant increase in the termination of pregnancy rate and a drop in perinatal mortality. There were, however, large differences in the detection rates between the two hospitals involved in this study, which shows that variation in skill and expertise can impact on performance and effectiveness of ultrasonography, and highlights the need for education, training, audit and quality control. This point is further emphasised by the low detection rate of major fetal malformations in the large RADIUS 1993 trial - only 17% of such babies were identified in the ultrasound screened group before 24 weeks of pregnancy. Based on the Helsinki 1990 trial results and other reports of observational data, this implies unsatisfactory diagnostic expertise. A combination of low detection rates of malformation, together with a gestational age limit of 24 weeks for legal termination of pregnancy in the RADIUS 1993 trial, produced minimal impact on perinatal mortality, unlike the Helsinki 1990 experience.

The majority of obstetric units in the developed world already practice routine early pregnancy ultrasonography. For those considering its introduction, the benefit of the demonstrated advantages needs to be considered against the theoretical possibility that the use of ultrasound during pregnancy could be harmful, and the need for additional resources. At present, there is no clear evidence that ultrasound examination during pregnancy is harmful. The findings from the follow-up of school children and teenagers, exposed as fetuses to ultrasound in the Norwegian and Swedish trials (Norway 1993; Sweden 1988) are generally reassuring; the finding that fewer children in the Norwegian ultrasound groups were right-handed was not confirmed by intention-to-treat analysis of long-term follow-up data from the Swedish trial. The Norwegian finding is difficult to interpret and may have been a chance observation that emanated from the large number of outcome measures assessed, or from the method of ascertainment. Alternatively, if it was a real consequence of ultrasound exposure, then it could imply that the effect of diagnostic ultrasound on the developing brain may alter developmental pathways. No firm conclusion can be reached from available data, and there is a need to study these children formally rather than to rely on a limited number of questionnaire responses obtained from the parents (Paneth 1998).

Financial costs also need to be considered. Calculations by the authors of the Radius report indicate that screening four million pregnant women in the USA at 200 dollars per scan would increase costs by one billion dollars per year (LeFevre 1993). While costs have been shown to be less in other countries (Henderson 2002; Roberts 2002), economic issues will still be relevant, particularly in low-resource settings. Clinicians, health planners, and pregnant women need to decide if these results justify the expense of providing routine ultrasound examination in early pregnancy. The early Helsinki 1990 data may have overestimated the efficiency of scans. Cost savings were assumed on the basis of decreased perinatal mortality, which was not borne out in other studies.

Maternal anxiety and satisfaction have not been well explored in the studies included in this review. Parents may not be fully informed about the purpose of routine ultrasonography and may be made anxious, or be inappropriately reassured by scans (Garcia 2002; Lalor 2007). Ultrasound scans are, however, popular - the potential enjoyment that parents can receive from seeing the image of their baby in utero is discussed elsewhere (Neilson 1995).

Overall completeness and applicability of evidence

The review includes several large trials, although the eligibility criteria of some trials (e.g. RADIUS 1993) mean that results may not be generalisable to all women.

The majority of studies were carried out in high-resource settings where overall levels of perinatal mortality are low and the contribution of major fetal abnormality to mortality is higher than in lower-resource settings. Findings in high-resource settings may not apply in less affluent settings.

Studies were carried out over three decades and technical advances in equipment, more widespread use of ultrasonography in the developed world, and training and expertise of operators are likely to have resulted in more effective sonography, particularly for the detection of fetal abnormalities. The two trials which evaluated detection of fetal abnormality are probably not relevant in today's setting.

Quality of the evidence

Overall, the review includes several large, well-designed trials but the lack of blinding is a problem common to all of the studies and this may have an effect on some outcomes. The quality of evidence was assessed for all review primary outcomes using GRADEpro 2014. The evidence for detection of fetal abnormality and multiple pregnancy before 24 weeks' gestation outcomes was judged as moderate. The quality of evidence was low for induction of labour for 'post-term pregnancy' and 'perinatal death' (all babies) outcomes. Downgrading of evidence was based on including studies with design limitations, imprecision of results and presence of heterogeneity.

Potential biases in the review process

The possibility of introducing bias was present at every stage of the reviewing process. We attempted to minimise bias in a number of ways; two review authors assessed eligibility for inclusion, carried out data extraction and assessed risk of bias. Each worked independently. Nevertheless, the process of assessing risk of bias, for example, is not an exact science and includes many personal judgements. Further, the process of reviewing research studies is known to be affected by prior beliefs and attitudes. It is difficult to control for this type of bias in the reviewing process.

While we attempted to be as inclusive as possible in the search strategy, the literature identified was predominantly written in English and published in North American and European journals. Although we did attempt to assess reporting bias, constraints of time meant that this assessment largely relied on information available in the published trial reports and thus, reporting bias was not usually apparent.

Agreements and disagreements with other studies or reviews

We have not been able to identify many other studies or reviews in this area. The few studies identified focus mainly on the identification of fetal anomalies in early pregnancy. One study (Drysdale 2002), is a pilot study for screening by ultrasound examination of all women, presenting to their community midwife before 12 weeks' gestation. The included women were at high risk and had abnormal pregnancies. The researchers found that early pregnancy ultrasound at 12 to 14 weeks' gestation was an effective method of identifying and screening for major abnormalities of pregnancy. Five of the seven major abnormalities diagnosed were identified before the 20-week anomaly scan. The authors emphasised the importance of using this in conjunction with an anomaly scan at around 20 weeks' gestation. Women found this method of screening acceptable. Rossi 2013, a systematic review, tested the accuracy of ultrasonography at 11 to 14 weeks of gestation for detection of fetal structural anomalies. Inclusion criteria were fetal anatomy examination performed at 11 to 14 weeks of gestation and confirmation of the early diagnosis of fetal malformations by postnatal or postmortem examination in addition to many others. In this systematic review, which included 19 studies, the overall fetal structural anomalies detection rate was 472 of 957 (51%). The authors stated that the number of fetal malformations remain undetected by early ultrasonography. We were unable to find other studies that address the other primary outcomes of this review.

AUTHORS' CONCLUSIONS

Implications for practice

Routine early pregnancy ultrasonography has been shown to detect multiple pregnancy earlier (*moderate quality evidence*), and to reduce induction of labour for post-term pregnancy (*low quality evidence*), both of which could be clinically useful if resources allow.

Implications for research

1. Other benefits which could result from better gestational age assessment, e.g. better management of pregnancies complicated by fetal growth retardation, need to be assessed in much larger studies than have been reported so far.

2. Earlier detection of twin pregnancies has not been translated into an improvement in fetal outcome. The continuing high perinatal mortality rate from multiple pregnancies is a legitimate cause for concern and requires study on a number of fronts; at least, early detection improves the potential for the proper scientific study of other, potentially useful, interventions in multiple pregnancies.

3. There is a need to look at the value of the detection of fetal abnormalities in terms of important outcomes (perinatal morbidity, mortality, burden of disease and long-term outcomes) in the context of newer imaging technologies and advancing skill in fetal abnormality detection.

4. In settings where resources are constrained, the question of the added value of ultrasonography should be revisited in the light of technological advances.

5. In the developed world, screening for Down's syndrome and fetal abnormality is moving into the first trimester and there is a need to evaluate the optimal timing of ultrasonography for gestational dating, multiple pregnancy detection including chorionicity determination, and fetal abnormality detection. There is also a need to determine whether one or more scans are needed to provide all this information in a way that is cost-effective and acceptable to women.

6. There is still some debate about the long-term outcomes associated with fetal exposure to ultrasound, especially multiple exposures and the use of colour Doppler in the first trimester.

7. There is a lack of evidence from trials on women's views; although ultrasound is perceived as a largely benign technology and is popular with women and their families, the finding of structural features in the fetus of uncertain significance can cause great upset. In the context of advancing technology, women's views should be explored as part of future evaluations.

ACKNOWLEDGEMENTS

R E F E R E N C E S

References to studies included in this review

Adelaide 1999 *{published data only}*

Crowther CA. Trial to assess whether ultrasound examination at the booking antenatal visit reduces the number of repeat screenings and results in earlier diagnosis of non-viable pregnancy/congenital abnormality. Personal communication 1992.

* Crowther CA, Kornman L, O'Callaghan S, George K, Furness M, Willson K. Is an ultrasound assessment of gestational age at the first antenatal visit of value? A randomised clinical trial. *British Journal of Obstetrics and Gynaecology* 1999;**106**(12):1273-9.

Alesund 1984 *{published data only}*

Eik-Nes SH. Effects of routine two-stage ultrasound screening in pregnancy: the Alesund randomised controlled trial revisited. Personal communication 1984.

Eik-Nes SH, Okland O. Ultrasound screening of pregnant women - a prospective randomized study. Diagnostic Ultrasound Imaging in Pregnancy. NIH Publication No.84-667. Washington: US Department of Health and Human Services, 1984:207-13.

* Eik-Nes SH, Okland O, Aure JC, Ulstein M. Ultrasound screening in pregnancy: a randomised controlled trial. *Lancet* 1984;**1**:1347.

Eik-Nes SH, Salvesen KA, Okland O, Vatten LJ. Routine ultrasound fetal examination in pregnancy: the 'Alesund' randomized controlled trial. *Ultrasound in Obstetrics & Gynecology* 2000;**15**(6):473-8.

Salvesen KA. Routine Ultrasonography in Utero and Development in Childhood - a Randomized Controlled Follow-up Study [MSc thesis]. University of Trondheim, 1993.

Helsinki 1990 *{published data only}*

Leivo T, Tuominen R, Saari-Kemppainen A, Ylostalo P, Karjalainen O, Heinonen OP. Cost-effectiveness of one-stage ultrasound screening in pregnancy: a report from the Helsinki ultrasound trial. *Ultrasound in Obstetrics & Gynecology* 1996;**7**(5):309-14.

Saari-Kemppainen A. Use of antenatal care services in a controlled ultrasound screening trial. *Acta Obstetricia et Gynecologica Scandinavica* 1995;**74**:12-4.

Saari-Kemppainen A, Karjalainen O, Ylostalo P. A randomised study of ultrasound screening during pregnancy. 12th FIGO World Congress of Gynecology and Obstetrics; 1988 October 23-28; Brazil. 1988:247-8.

Saari-Kemppainen A, Karjalainen O, Ylostalo P. Ultrasound screening and perinatal mortality: a controlled clinical trial. Proceedings of 12th European Congress of Perinatal Medicine; 1990 Sept 11-14; Lyon, France. 1990:36.

Saari-Kemppainen A, Karjalainen O, Ylostalo P, Heinonen OP. Fetal anomalies in a controlled one-stage ultrasound screening trial. A report from the Helsinki Ultrasound Trial. *Journal of Perinatal Medicine* 1994;**22**:279-89.

* Saari-Kemppainen A, Karjalainen O, Ylostalo P, Heinonen OP. Ultrasound screening and perinatal mortality: controlled trial of systematic one-stage screening in pregnancy. Lancet 1990; Vol. 336:387-91.

Johannesburg 2007 *{published data only}*

* van Dyk B, Motto JA, Buchmann EJ. Routine second-trimester ultrasound for low risk pregnancies in a South African community. *International Journal of Gynecology & Obstetrics* 2007;**98**(3):257-8.

van Dyk B, Motto JA, Buchmann EJ. The value of routine mid-trimester ultrasound in low-risk pregnancies at primary care level. *Health SA Gesondheid* 2008;**13**(4):41-9.

London 1982 *{published data only}*

Bennett MJ, Little G, Dewhurst J, Chamberlain GVP. Predictive value of ultrasound measurement in early pregnancy: a randomized controlled trial. *British Journal of Obstetrics and Gynaecology* 1982;**89**:338-41.

Missouri 1990 *{published data only}*

Ewigman B, LeFevre M, Hesser J. A randomized trial of routine prenatal ultrasound. *Obstetrics & Gynecology* 1990;**76**:189-94.

Norway 1993 *{published data only}*

Salvesen KA. Routine ultrasonography in utero and development in childhood - a randomized controlled follow-up study. *Acta Obstetricia et Gynecologica Scandinavica* 1995;**74**:166-7.

Salvesen KA, Bakketeig LS, Eik-Nes SH, Undheim JO, Okland O. Routine ultrasonography in utero and school performance at age 8-9 years. *Lancet* 1992;**339**:85-9.

Salvesen KA, Jacobsen G, Vatten LJ, Eik-Nes SH, Bakketeig LS. Routine ultrasonography in utero and subsequent growth during childhood. *Ultrasound in Obstetrics & Gynecology* 1993;**3**:6-10.

* Salvesen KA, Vatten LJ, Eik-Nes SH, Hugdahl K, Bakketeig LS. Routine ultrasonography in utero and subsequent handedness and neurological development. *BMJ* 1993;**307**:159-64.

Salvesen KA, Vatten LJ, Jacobsen G, Eik-Nes SH, Okland O, Molne K, et al. Routine ultrasonography in utero and subsequent vision and hearing at primary school age. *Ultrasound in Obstetrics & Gynecology* 1992;**2**:243-7.

Oxford 2006 *{published data only}*

* Harrington DJ, MacKenzie IZ, Thompson K, Fleminger M, Greenwood C. Does a first trimester dating scan using crown rump length measurement reduce the rate of induction of labour for prolonged pregnancy? An uncompleted randomised controlled trial of 463 women. *BJOG: an international journal of obstetrics and gynaecology* 2006;**113**(2):171-6.

Harrington DJ, Mackenzie IZ, Chamberlain P, Greenwood C. Does a first trimester crown-rump length (CRL) measurement reduce the rate of elective timed delivery for post dates? A

randomised control trial [abstract]. *Journal of Obstetrics and Gynaecology* 2004;**24**(Suppl 1):S22.

MacKenzie I. The effect of first trimester crown rump length (CRL) measurement rates of labour for postdates. Research Findings Register www.refer.nhs.uk (accessed 7 March 2006).

RADIUS 1993 *{published data only}*

Crane JP, LeFevre ML, Winborn RC, Evans JK, Ewigman G, Bain RP, et al. A randomized trial of prenatal ultrasonographic screening: impact on the detection, management, and outcome of anomalous fetuses. *American Journal of Obstetrics and Gynecology* 1994;**171**:392-9.

* Ewigman BG, Crane JP, Frigoletto FD, LeFevre ML, Bain RP, McNellis D, et al. Effect of prenatal ultrasound screening on perinatal outcome. *New England Journal of Medicine* 1993;**329**:821-7.

Frigoletto FD Jr, Ewigman BG, Crane JP, LeFevre ML, Bain RP, McNellis D. Routine ultrasound screening for all pregnant women: does it make a difference?. *Acta Obstetrica Et Gynaecologica Japonica* 1997;**49**:452.

Harlow BL, Frigoletto FD, Cramer DW, Evans JK, Bain RP, Ewigman B, et al. Epidemiologic predictors of Cesarean section in nulliparous patients at low risk. *American Journal of Obstetrics and Gynecology* 1995;**172**:156-62.

LeFevre ML, Bain RP, Ewigman BG, Frigoletto FD, Crane JP, McNellis D, et al. A randomized trial of prenatal ultrasonographic screening: impact on maternal management and outcome. *American Journal of Obstetrics and Gynecology* 1993;**169**:483-9.

LeFevre ML, Evans JK, Ewigman B. Is smoking an indication for prenatal ultrasonography? RADIUS Study Group;. Archives of Family Medicine 1995; Vol. 4:120-3.

Sweden 1988 *{published data only}*

Axelsson O. Estimation of gestational age by measurement of the biparietal diameter in the second trimester: preliminary results of a randomized trial. Proceedings of 6th Congress of the European Federation of Societies for Ultrasound in Medicine and Biology; 1987 June 14-18; Helsinki, Finland. 1987.

Kieler H, Ahlsten G, Haglund B, Salvesen K, Axelsson O. Routine ultrasound screening in pregnancy and the children's subsequent neurologic development. *Obstetrics & Gynecology* 1998;**91**(5 Pt 1):750-6.

Kieler H, Axelsson O, Haglund B, Nilsson S, Salvesen KA. Routine ultrasound screening in pregnancy and the children's subsequent handedness. *Early Human Development* 1998;**50**(2):233-45.

Kieler H, Haglund B, Waldenstrom U, Axelsson O. Routine ultrasound screening in pregnancy and the children's subsequent growth, vision and hearing. *British Journal of Obstetrics and Gynaecology* 1997;**104**(11):1267-72.

Stalberg K, Axelsson O, Haglund B, Hultman CM, Lambe M, Kieler H. Prenatal ultrasound exposure and school achievement

in teenagers; follow-up of a randomized controlled trial. *Ultrasound in Obstetrics and Gynecology* 2008;**32**(3):306.

Stålberg K, Axelsson O, Haglund B, Hultman CM, Lambe M. Prenatal ultrasound exposure and children's school performance at age 15-16: follow-up of a randomized controlled trial. *Ultrasound in Obstetrics and Gynecology* 2009;**34**:297-303.

Waldenstrom U, Axelsson O, Nilsson S. Ultrasonic dating of pregnancies: effect on incidence of SGA diagnoses. A randomised controlled trial. Early Human Development 1992; Vol. 30:75-9.

* Waldenstrom U, Axelsson O, Nilsson S, Eklund G, Fall O, Lindeberg S, et al. Effects of routine one-stage ultrasound screening in pregnancy: a randomised controlled trial. *Lancet* 1988;**2**:585-8.

Trondheim 1984 *{published data only}*

Bakketeig LS, Jacobsen G, Brodtkorb CJ, Eriksen BC, Eik-Nes SH, Ulstein MK, et al. Randomised controlled trial of ultrasonographic screening in pregnancy. *Lancet* 1984;**2**:207-10.

Tygerberg 1996 *{published data only}*

Geerts L, Brand E, Theron GB. Routine obstetric ultrasound examinations in South Africa: Cost and effect on perinatal outcome - A prospective randomised controlled trial. Proceedings of the 14th Conference on Priorities in Perinatal Care in South Africa; 1995 March 7-10; South Africa. 1995:130-3.

* Geerts LTGM, Brand EJ, Theron GB. Routine obstetric ultrasound examinations in South Africa: cost and effect on perinatal outcome - a prospective randomised controlled trial. *British Journal of Obstetrics and Gynaecology* 1996;**103**:501-7.

References to studies excluded from this review

Belanger 1996 *{published data only}*

Belanger K, Hobbins JC, Muller JP, Howard S. Neurological testing in ultrasound exposed infants. *American Journal of Obstetrics and Gynecology* 1996;**174**(1 Pt 2):413.

Bennett 2004 *{published data only}*

Bennett K, Crane J, O'Shea P, Lacelle J, Hutchens D, Copel J. Combined first and second trimester ultrasound screening is effective in reducing postterm labor induction rates: a randomized controlled trial [abstract]. *American Journal of Obstetrics and Gynecology* 2002;**187**(6 Pt 2):S68.

* Bennett KA, Crane JMG, O'Shea P, Lacelle J, Hutchens D, Copel JA. First trimester ultrasound screening is effective in reducing postterm labor induction rates: a randomized controlled trial. *American Journal of Obstetrics and Gynecology* 2004;**190**:1077-81.

Duff 1993 *{published data only}*

Duff G. A randomised controlled trial in a hospital population of ultrasound measurement screening for the small for dates baby. Proceedings of 2nd International Scientific Meeting of the Royal College of Obstetricians and Gynaecologists; 1993 Sept 7-10; Hong Kong. 1993:90.

* Duff GB. A randomized controlled trial in a hospital population of ultrasound measurement screening for the small for dates baby. *Australian and New Zealand Journal of Obstetrics and Gynaecology* 1993;**33**:374-8.

Georgsson Ohman 2010 *{published data only}*

Georgsson Ohman S, Waldenstrom U, Georgsson Ohman S, Waldenstrom U. Effect of first-trimester ultrasound screening for Down syndrome on maternal-fetal attachment--a randomized controlled trial. *Sexual & Reproductive Healthcare* 2010;**1**(3):85-90.

Hong Kong *{published data only}*

Chen M, Lee CP, Lam YH, Tang RY, Chan BC, Wong SF, et al. Comparison of nuchal and detailed morphology ultrasound examinations in early pregnancy for fetal structural abnormality screening: a randomized controlled trial. *Ultrasound in Obstetrics and Gynecology* 2008;**31**(2):136-46.

Larsen 1992 *{published data only}*

Larsen T, Larsen JK, Petersen S, Greisen G. Detection of small-for-gestational-age fetuses by ultrasound screening in a high risk population: a randomized controlled study. *British Journal of Obstetrics and Gynaecology* 1992;**99**:469-74.

Leung 2006 *{published data only}*

Leung KY, Ngai CS, Lee A, Chan HY, Leung WC, Lee CP, et al. The effects on maternal anxiety of two-dimensional versus two-plus three-/four-dimensional ultrasound in pregnancies at risk of fetal abnormalities: A randomized study. *Ultrasound in Obstetrics and Gynecology* 2006;**28**(3):249-54.

Owen 1994 *{published data only}*

Owen P, Donnet L, Ogston S, Christie A, Patel N, Howie P. A study of fetal growth velocity. *British Journal of Obstetrics and Gynaecology* 1994;**101**:270.

Rustico 2005 *{published data only}*

Righetti PL, Dell'Avanzo M, Grigio M, Nicolini U. Maternal/paternal antenatal attachment and fourth-dimensional ultrasound technique: a preliminary report. *British Journal of Psychology* 2005;**96**(Pt 1):129-37.

* Rustico MA, Mastromatteo C, Grigio M, Maggioni C, Gregori D, Nicolini U. Two-dimensional vs. two- plus four-dimensional ultrasound in pregnancy and the effect on maternal emotional status: a randomized study. *Ultrasound in Obstetrics & Gynecology* 2005;**25**:468-72.

Saltvedt 2006 *{published data only}*

Georgsson Ohman S, Saltvedt S, Grunewald C, Waldenstrom U. Does fetal screening affect the women's worries about the health of their baby? A randomized controlled trial of ultrasound screening for Down's syndrome versus routine ultrasound screening. *Acta Obstetricia et Gynecologica Scandinavica* 2004;**83**:634-40.

Saltvedt S, Almstrom H, Kublick M, Reilly M, Valentin L, Grunewald C. Ultrasound dating at 12-14 or 15-20 weeks of gestation? A prospective cross-validation of established dating formulae in a population of in-vitro fertilized pregnancies

randomized to early or late dating scan. *Ultrasound in Obstetrics & Gynecology* 2004;**24**:42-50.

Saltvedt S, Almstrom H, Kublickas M, Valentin L, Bottinga R, Bui TH, et al. Screening for Down syndrome based on maternal age or fetal nuchal translucency: a randomized controlled trial in 39 572 pregnancies. *Ultrasound in Obstetrics and Gynecology* 2005;**25**:537-45.

* Saltvedt S, Almstrom H, Kublickas M, Valentin L, Grunewald C. Detection of malformations in chromosomally normal fetuses by routine ultrasound at 12 or 18 weeks of gestation - a randomised controlled trial in 39,572 pregnancies. *BJOG: an international journal of obstetrics and gynaecology* 2006;**113**(6):664-74.

Westin M, Saltvedt S, Bergman G, Kublickas M, Almstrom H, Grunewald C, et al. Routine ultrasound examination at 12 or 18 gestational weeks for prenatal detection of major congenital heart malformations? A randomised controlled trial comprising 36,299 fetuses. *BJOG: an international journal of obstetrics and gynaecology* 2006;**113**(6):675-82.

Schifano 2010 *{published data only}*

Schifano M, Luchi C, Sceusa F, Nanini C, Pepe A, Mannella P, et al. Women's attitude towards ultrasound scanning in the first trimester of pregnancy: two-dimensional versus two-plus four-dimensional ultrasound effects on maternal emotional status. *Journal of Psychosomatic Obstetrics and Gynecology* 2010;**31**(s1):65.

Schwarzler 1999 *{published data only}*

Schwarzler P, Senat MV, Holden D, Bernard JP, Masroor T, Ville Y. Feasibility of the second-trimester ultrasound examination in an unselected population at 18, 20 or 22 weeks of pregnancy: a randomised trial. *Ultrasound in Obstetrics & Gynecology* 1999;**14**:92-7.

Votino 2012 *{published data only}*

Votino C, Kacem Y, Dobrescu O, Dessy H, Cos T, Foulon W, et al. Use of a high-frequency linear transducer and MTI filtered color flow mapping in the assessment of fetal heart anatomy at the routine 11 to 13 + 6-week scan: a randomized trial. *Ultrasound in Obstetrics & Gynecology* 2012;**39**(2):145-51.

Wald 1988 *{unpublished data only}*

Wald NJ. Randomised controlled trial of routine dating ultrasound in pregnancy. Personal communication 1988.

References to studies awaiting assessment

Newcastle 2004 *{published data only}*

Deverill M, Snaith V, Howel D, Hewison J, Sturgiss S, Robson S. The Newcastle randomised controlled trial of early screening for fetal abnormality - women's preferences for early information on fetal status and cost-effectiveness analysis [abstract]. *Journal of Obstetrics and Gynaecology* 2004;**24**(Suppl 1):S20.

* Snaith V, Howel D, Deverill M, Hewison J, Sturgiss S, Robson S. The Newcastle randomised controlled trial of early ultrasound screening for fetal abnormality (FA) - termination of pregnancy

for FA and psychological consequences [abstract]. *Journal of Obstetrics and Gynaecology* 2004;**24**(Suppl 1):S19.

Snaith VJ, Howel D, Chadwick T, Deverill M, Hewison J, Sturgiss SN, et al. First trimester ultrasound screening - the psychological consequences of termination of pregnancy for foetal abnormality. *Journal of Reproductive and Infant Psychology* 2004;**22**(3):239.

Sturgiss S, Howel D, Snaith V, Deverill M, Hewison J, Robson S. The Newcastle randomized controlled trial of early ultrasound screening for fetal abnormality: impact on gestation at diagnosis [abstract]. *Journal of Obstetrics and Gynaecology* 2005;**25 Suppl 1**:S20.

References to ongoing studies

McClure 2014 *{published data only}*

McClure EM, Nathan RO, Saleem S, Esamai F, Garces A, Chomba E, et al. First look: a cluster-randomized trial of ultrasound to improve pregnancy outcomes in low income country settings. BMC Pregnancy and Childbirth. NCT01990625 2014; Vol. 14, issue 1:73.

Additional references

Barrett 1991

Barrett J, Brinson J. Evaluation of obstetric ultrasonography at the first prenatal visit. *American Journal of Obstetrics and Gynecology* 1991;**165**:1002-5.

Berghella 2013

Berghella V, Baxter JK, Hendrix NW. Cervical assessment by ultrasound for preventing preterm delivery. *Cochrane Database of Systematic Reviews* 2013, Issue 1. [DOI: 10.1002/14651858.CD007235.pub3]

Bricker 2000

Bricker L, Garcia J, Henderson J, Mugford M, Neilson J, Roberts T, et al. Ultrasound screening in pregnancy: a systematic review of the clinical effectiveness, cost-effectiveness and women's views. *Health Technology Assessment* 2000;**4**:1-193.

Bricker 2008

Bricker L, Neilson JP, Dowswell T. Routine ultrasound in late pregnancy (after 24 weeks' gestation). *Cochrane Database of Systematic Reviews* 2008, Issue 4. [DOI: 10.1002/14651858.CD001451.pub3]

Crane 2008

Crane JM, Hutchens D, Crane JMG. Transvaginal sonographic measurement of cervical length to predict preterm birth in asymptomatic women at increased risk: a systematic review. *Ultrasound in Obstetrics & Gynecology* 2008;**31**(5):579-87.

Dodd 2005

Dodd JM, Crowther CA. Evidence-based care of women with a multiple pregnancy. *Best Practice & Research. Clinical Obstetrics & Gynaecology* 2005;**19**:131-53.

Drysdale 2002

Drysdale K, Ridley D, Walker K, Higgins B, Dean T. First-trimester pregnancy scanning as a screening tool for high-risk and abnormal pregnancies in a district general hospital setting. *Journal of Obstetrics and Gynaecology* 2002;**22**(2):159-65.

EFSUMB 1995

European Federation of Societies for Ultrasound in Medicine and Biology. Watchdog Committee, 1994 clinical safety statement. *European Journal of Ultrasound* 1995;**2**:77.

Garcia 2002

Garcia J, Bricker L, Henderson J, Martin MA, Mugford M, Neilson J, et al. Women's views of pregnancy ultrasound: a systematic review. *Birth* 2002;**29**(4):225-50.

Geirsson 1991

Geirsson R, Busby-Earle R. Certain dates may not provide a reliable estimate of gestational age. *British Journal of Obstetrics and Gynaecology* 1991;**98**:108-9.

GRADEpro 2014 [Computer program]

McMaster University. GRADEpro. [Computer program on www.gradepro.org]. Version 2015. McMaster University, 2014.

Gulmezoglu 2012

Gulmezoglu AM, Crowther CA, Middleton P, Heatley E. Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database of Systematic Reviews* 2012, Issue 6. [DOI: 10.1002/14651858.CD004945.pub3]

Henderson 2002

Henderson J, Bricker L, Roberts T, Mugford M, Garcia J, Neilson J. British National Health Service's and women's costs of antenatal ultrasound screening and follow-up tests. *Ultrasound in Obstetrics & Gynecology* 2002;**20**(2):154-62.

Higgins 2011

Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Lalor 2007

Lalor JG, Devane D. Information, knowledge and expectations of the routine ultrasound scan. *Midwifery* 2007;**23**(1):13-22.

Lee 2006

Lee YM, Cleary-Goldman J, Thaker HM, Simpson LL. Antenatal sonographic prediction of twin chorionicity. *American Journal of Obstetrics and Gynecology* 2006;**195**:863-7. [MEDLINE: 16949427]

LeFevre 1993

LeFevre ML, Bain RP, Ewigman BG, Frigoletto FD, Crane JP, McNellis D, et al. A randomized trial of prenatal ultrasonographic screening: impact on maternal management and outcome. *American Journal of Obstetrics and Gynecology* 1993;**169**:483-9.

Leivo 1996

Leivo T, Tuominen R, Saari Kemppainen A, Ylostalo P, Karjalainen O, Heinonen OP. Cost-effectiveness of one-stage ultrasound screening in pregnancy: a report from the Helsinki ultrasound trial. *Ultrasound in Obstetrics & Gynecology* 1996;**7**(5):309-14.

Neilson 1995

Neilson JP. High vs low feedback to mother at fetal ultrasound. [revised 12 May 1994]. In: Enkin MW, Keirse MJNC, Neilson JP, Crowther CA (eds) Pregnancy and Childbirth Module. In: The Cochrane Pregnancy and Childbirth Database [database on disk and CDROM] The Cochrane Collaboration; Issue 2, Oxford: Update Software; 1995.

Owen 1997

Owen J, Wenstrom K. The effect of inaccurate gestational age estimation on the multiple marker screening test (MMST) for fetal Down Syndrome. *American Journal of Obstetrics and Gynecology* 1997;**176**:A297.

Paneth 1998

Paneth N. Prenatal sonography - safe or sinister?. *Lancet* 1998;**352**:5-6.

Peek 1994

Peek M, Devonald K, Beilby R, Ellwood D. The value of routine early pregnancy ultrasound in the antenatal booking clinic. *Australian and New Zealand Journal of Obstetrics and Gynaecology* 1994;**34**:140-3.

Persson 1983

Persson PH, Kullander S. Long-term experience of general ultrasound screening in pregnancy. *American Journal of Obstetrics and Gynecology* 1983;**146**:942-7.

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Roberts 2002

Roberts T, Henderson J, Mugford M, Bricker L, Neilson J, Garcia J, et al. Antenatal ultrasound screening for fetal abnormalities: a systematic review of studies of cost and cost effectiveness. *BJOG: an international journal of obstetrics and gynaecology* 2002;**109**(1):44-56.

Rossi 2013

Rossi AC, Prefumo F. Accuracy of ultrasonography at 11-14 weeks of gestation for detection of fetal structural anomalies: a systematic review. *Obstetrics and Gynecology* 2013;**122**(6):1160-7.

Saari-Kemppainen 1990

Saari-Kemppainen A, Karjalainen O, Ylostalo P, Heinonen OP. Ultrasound screening and perinatal mortality: controlled trial of systematic one-stage screening in pregnancy. *Lancet* 1990;**336**:387-91.

Schunemann 2009

Schunemann HJ. GRADE: from grading the evidence to developing recommendations. A description of the system and a proposal regarding the transferability of the results of clinical research to clinical practice [GRADE: Von der Evidenz zur Empfehlung. Beschreibung des Systems und Losungsbeitrag zur Ubertragbarkeit von Studienergebnissen]. *Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen* 2009;**103**(6):391-400. [PUBMED: 19839216]

References to other published versions of this review

Whitworth 2010

Whitworth M, Bricker L, Neilson JP, Dowswell T. Ultrasound for fetal assessment in early pregnancy. *Cochrane Database of Systematic Reviews* 2010, Issue 4. [DOI: 10.1002/14651858.CD007058.pub2]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

Adelaide 1999

Methods	RCT (women randomised).
Participants	Setting: Adelaide, Australia, 2 hospitals. Study carried out 1991-5.
	648 women attending hospital for their first antenatal visit.
	Inclusion criteria: women attending for antenatal care < 17 weeks' gestation, no previous ultrasound scans this pregnancy, expected to deliver at the study hospital, no clear indication for ultrasound at the first visit.
Interventions	Intervention group (n = 321): ultrasound scan at 11-14 weeks' gestation (at the antenatal booking visit) performed by medical staff with ultrasound training or in the ultrasound department.
	Control group (n = 327): routine care.

Adelaide 1999 *(Continued)*

All women were offered routine scans at 18-20 weeks and completed assessments of anxiety and feelings about pregnancy.

Outcomes	Number of women whose EDD was adjusted by 10 or more days at the 18-20 weeks' scan. Number of women booked for fetal morphology scan at sub-optimal gestation (< 17 or > 20 weeks). Number of women needing repeat serum screening. Number of women who felt worried about pregnancy at the end of the first visit.
Notes	Numbers of twins reported varies between table 1 and table 2 and table 7.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number tables. Randomisation carried out by external university clinical trials unit.
Allocation concealment (selection bias)	Low risk	Consecutively numbered opaque, sealed envelopes. Envelopes prepared by researcher not involved in clinical care.
Blinding (performance bias and detection bias) Women	High risk	Women would have been aware of group assignment due to the nature of the intervention.
Blinding (performance bias and detection bias) Clinical staff	Unclear risk	Reports not available to staff carrying out 18-20 week scans.
Blinding (performance bias and detection bias) Outcome assessors	Unclear risk	Staff carrying out 18-20 week scan were masked to group assignment, but women would have been aware of their group assignment when filling out questionnaire at first visit.
Incomplete outcome data (attrition bias) All outcomes	Low risk	648 randomised (321 in intervention group and 327 controls). All women randomised were included in the analysis for the primary outcome. 37 women miscarried before the mid-trimester morphology scan (18 in the intervention group and 19 in the control group). Few women were lost to follow-up (17). Pregnancy outcome data were available for 296/321 of the intervention group and 298/327 of the control group. There were missing data (< 10%) for some outcomes.
Selective reporting (reporting bias)	Low risk	Prespecified outcomes are reported.
Other bias	Low risk	Groups balanced at baseline. Primary analysis by ITT.

Alesund 1984

Methods	RCT. Individual randomisation. Alesund, Norway.
Participants	1628 women attending their first antenatal visit at the clinics of 35 general practitioners (nearly all women in that geographical area, including those with 'high-risk' pregnancies). Recruitment 1979-1981.
Interventions	Intervention group: routine ultrasound examinations at 18 weeks (biparietal diameter measured to predict EDD, multiple pregnancies noted, placenta located and general examination of the fetus) and 32 weeks (biparietal diameter and abdominal circumference, final placental location and presentation

Alesund 1984 *(Continued)*

noted) with additional examination at 36 weeks if fetus small-for-gestational age or presenting by the breech.

Control group: selective examination for specific clinical indications only (77% of women in the control group did not have an ultrasound examination).

Ultrasound examinations performed by 1 of 2 experienced doctors.

Outcomes	Primary outcome: induction of labour for 'post-term' pregnancy; secondary outcomes: indices of perinatal mortality and morbidity.
Notes	The data that have been entered into this version of the review are derived from only those pregnancies that were singleton, except for perinatal mortality rates which are calculated from all pregnancies. There were 10 multiple pregnancies in the screened group and 76 in the control group.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Described as the "sealed envelope method."
Blinding (performance bias and detection bias) Women	High risk	Blinding not attempted.
Blinding (performance bias and detection bias) Clinical staff	High risk	Blinding not attempted.
Blinding (performance bias and detection bias) Outcome assessors	High risk	No blinding.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Very low levels of missing data (although denominators were not provided for all outcomes reported). Some discrepancies between the 1984 and 2000 publications.
Selective reporting (reporting bias)	Low risk	Main outcome, and power calculation for post-term induced labour. No selective reporting bias apparent.
Other bias	Unclear risk	Some baseline imbalance between groups. 64% of the intervention group compared to 69% of the control group were non-smokers (P = 0.02).

Helsinki 1990

Methods	RCT (women randomised). Finland.
Participants	All women attending 1 of 64 health centres. Recruitment 1986-1987.
	9310 women randomised (women were included even if they had had a previous scan elsewhere).
Interventions	Intervention group: (4691 randomised); ultrasound examination at 16-20 weeks.

Helsinki 1990 *(Continued)*

	Control group: routine care (selective scanning for specific reasons).
Outcomes	Fetal outcome and clinical interventions.
Notes	77% of women in the control group underwent at least 1 ultrasound scan during pregnancy. Mean scans per pregnancy: 2.1 (study group), 1.8 (control group).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence genera- tion (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Described as "by sealed envelope method".
Blinding (performance bias and detection bias) Women	High risk	Blinding not attempted.
Blinding (performance bias and detection bias) Clinical staff	High risk	No blinding.
Blinding (performance bias and detection bias) Outcome assessors	High risk	No blinding.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Small loss to follow-up (< 1%). Those women that did not attend for screening were included in the analysis as part of the intervention group. In the intervention group of 4691 randomised 9 women were not pregnant, there were 265 miscarriages before screening and 26 women had pregnan- cy terminations before screening, a further 6 women were found to be not pregnant after screening. In the control group of 4619 women randomised, 2 women were not pregnant, 284 had spontaneous miscarriage before screen- ing, and 21 had pregnancy terminations before screening.These women have not been included in the denominators in the analyses in this review.
Selective reporting (re- porting bias)	Low risk	None apparent.
Other bias	Low risk	None noted.

Johannesburg 2007

Methods	Quasi-randomised trial in South Africa.
Participants	Setting: semi-urban health district serving a working class population in what was described as a "re- source-constrained setting". 955 women enrolled (804 women available for analysis). Low-risk women. Twin pregnancies were in- cluded (6 intervention and 2 control).
Interventions	Intervention group: ultrasound screening at between 18 to 23 weeks.

Johannesburg 2007 *(Continued)*

	Control group: no routine ultrasound.
Outcomes	Inductions for post-term pregnancy, miscarriage and perinatal mortality, fetal and neonatal outcomes and clinical interventions.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Allocation by day of clinic attendance.
Allocation concealment (selection bias)	High risk	Group allocation could be anticipated in advance.
Blinding (performance bias and detection bias) Women	High risk	Not feasible.
Blinding (performance bias and detection bias) Clinical staff	High risk	Blinding not attempted.
Blinding (performance bias and detection bias) Outcome assessors	High risk	No blinding.
Incomplete outcome data (attrition bias) All outcomes	High risk	955 women were enrolled (it was not clear how many were randomised to each group). 151 women were lost to follow-up (15.7%) leaving 804 women for analysis. There were further missing data for some outcomes. Reasons for attrition were not described, and it was not clear that loss was balanced across groups.
Selective reporting (reporting bias)	Unclear risk	Relevant outcomes have been reported, but they were not prespecified in the paper.
Other bias	Unclear risk	It was reported that there was no baseline imbalance between groups, but characteristics of women in the intervention and control groups were not described.

London 1982

Methods	Quasi-randomised trial.
Participants	1062 women attending 3 obstetric clinics in a maternity hospital in London, UK.
Interventions	All women had an ultrasound at 16 weeks. BPD measured.

Intervention group: results of scan were recorded in patient notes. The BPD was used to calculate an EDD and the estimated EDD (from menstrual history) was revised if scan revealed 2 weeks or more difference in estimated gestational age. |

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

London 1982 *(Continued)*

Control group: results of scan were not revealed (if a multiple pregnancy had not been identified by 26 weeks' gestation then the code was broken and this was revealed). 30% (161/531) in the comparison group had results revealed due to clinician concerns.

Outcomes	Muliple pregnancy, EDD, perinatal mortality, and birthweight.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quasi-randomisation. Women "were divided into two groups according to the last digit of the hospital number".
Allocation concealment (selection bias)	High risk	Group allocation could be anticipated in advance of recruitment.
Blinding (performance bias and detection bias) Women	Unclear risk	Women would not have been aware of group assignment because all women received an ultrasound.
Blinding (performance bias and detection bias) Clinical staff	High risk	All women received an ultrasound, so technicians would not have been aware of allocation. Clinical staff would have been aware of status because results for the control group were not disclosed.
Blinding (performance bias and detection bias) Outcome assessors	High risk	Results of the ultrasound were not disclosed for women in the control group, but the absence of the report would be apparent to clinicians and outcome assessors.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Low levels of missing data. 1095 women were randomised, 4 women were lost to follow-up, there were 13 early miscarriages and 4 fetal deaths before 16 weeks, 4 women were not pregnant (it was not clear which group these women were in or whether loss was balanced across groups).
Selective reporting (reporting bias)	Unclear risk	Main outcomes were gestational age at delivery and induction but many other outcomes were reported - it was not clear whether they had been prespecified in the study protocol.
Other bias	Unclear risk	Women with multiple pregnancies not diagnosed by 26 weeks' gestation in the control group had the diagnosis revealed.

Missouri 1990

Methods	RCT (individual randomisation).
Participants	915 women attending for first prenatal visit at 8 general physician and 8 obstetrician clinics in Missouri, USA (study carried out 1984-1986). Exclusion criteria: women were excluded if an ultrasound was indicated at recruitment.
Interventions	Intervention group (n = 459): ultrasound scan at 10-18 weeks (most carried out between 10-12 weeks) to estimate gestational age and detect multiple pregnancy (fetal viability and uterine abnormalities noted). Control group (n = 456): usual care (scan when indicated).

Missouri 1990 *(Continued)*

Outcomes	Induction for post-dates pregnancy, detection of multiple pregnancy before 24 weeks, adverse perinatal outcome (death, admission to NICU for more than 3 days, 5 mins Apgar score < 6).
Notes	Exclusion criteria were such that 58% of those approached were not eligible for inclusion. 24% of the usual care group received scans but were analysed according to randomisation group.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as "randomization cards" in blocks of 4 and stratified by practice site.
Allocation concealment (selection bias)	Low risk	"sequentially numbered, sealed, opaque envelopes." Described as "double consent randomization", patients consented after they were aware of their randomisation group but were asked whether their records could be used in the analysis, "patients consenting to the use of their records but refusing to receive ultrasound were retained in the original assigned group for the purposes of analysis".
Blinding (performance bias and detection bias) Women	Unclear risk	9% lost to follow-up after randomisation.
Blinding (performance bias and detection bias) Clinical staff	High risk	Clinicians caring for the women were aware whether the women had received an ultrasound scan or not.
Blinding (performance bias and detection bias) Outcome assessors	High risk	Research staff collecting outcome data aware of whether women had received a scan or not. No comment with regards to staff analysing the data.
Incomplete outcome data (attrition bias) All outcomes	High risk	(Only 42% of women screened were eligible for inclusion in this trial). Of 415 randomised to the intervention group 11 refused consent and 46 were lost to follow-up. Of the 456 randomised to the control group 5 refused and 38 were lost to follow-up. A further 25 women in the intervention group refused scans but were included in the analyses according to randomisation group.
Selective reporting (reporting bias)	Low risk	Main outcomes gestational age and diagnosis of multiple pregnancies.
Other bias	Low risk	

Norway 1993

Methods	Long-term follow-up of 2 RCTs (Alesund 1984; Trondheim 1984) carried out 1979-81 in Norway.
Participants	*See* Alesund 1984 and Trondheim 1984. 2428 children (singletons) followed up at 8-9 years of age.
Interventions	Mothers of children in the intervention group had been offered routine ultrasound at 18 and 32 weeks' gestation. The comparison group had ultrasound selectively (by clinical indication).
Outcomes	Follow-up at 8-9 years, neurological, behavioural, and developmental outcomes.

Norway 1993 *(Continued)*

Notes	It was not clear whether long-term developmental outcomes had been specified at the outset.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	*See* Alesund 1984 and Trondheim 1984 studies. Not described.
Allocation concealment (selection bias)	Unclear risk	Described as "sealed envelope method".
Blinding (performance bias and detection bias) Women	High risk	Blinding not attempted.
Blinding (performance bias and detection bias) Clinical staff	High risk	Blinding not attempted
Blinding (performance bias and detection bias) Outcome assessors	Unclear risk	No details given
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Small loss to follow-up for the outcomes assessed in labour or the early post-natal period. For some longer-term outcomes there were higher levels of missing data (e.g. for outcomes on handedness complete data were available for 69% of the original sample).
Selective reporting (reporting bias)	Unclear risk	It was not clear whether the longer-term outcomes had been specified in the original study protocols.
Other bias	Unclear risk	There were differences between the Trondheim and Alesund samples which makes interpretation of the results difficult. The Alesund sample included most pregnant women, whereas the Trondheim sample was included more low-risk pregnancies.

Oxford 2006

Methods	RCT with block randomisation.
Participants	Setting: Oxford UK. 463 women attending 20 GP practices for pregnancy confirmation. Inclusion criteria: women attending in the first trimester with no obstetric indication for a first trimester scan.
Interventions	Intervention group: first trimester ultrasound scan (8-12 weeks) measuring crown-rump length by senior sonographer. EDD calculated if there were more than 5 days difference from menstrual dates. Control group: no first trimester scan. EDD from menstrual dates. Both groups had a routine anomaly scan at 18-20 weeks.
Outcomes	Induction of labour for post-term pregnancy.

Oxford 2006 *(Continued)*

Notes	Trial stopped early.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Block randomisation (block size 6).
Allocation concealment (selection bias)	Low risk	Consecutively numbered, opaque, sealed envelopes.
Blinding (performance bias and detection bias) Women	High risk	Not attempted
Blinding (performance bias and detection bias) Clinical staff	High risk	Results of scans in case notes.
Blinding (performance bias and detection bias) Outcome assessors	High risk	Results of scans were in case notes.
Incomplete outcome data (attrition bias) All outcomes	Low risk	463 randomised, 393 available for follow-up after 24 weeks. 4 women in each group were lost to follow-up. Report that the analysis was by ITT. (Loss before 24 weeks included miscarriage (24 in the intervention and 29 in the control group) these women were included in the ITT analysis.)
Selective reporting (reporting bias)	Low risk	The study reported on the pre-defined outcome measures
Other bias	High risk	The study was not completed. Private first trimester scans were introduced while the study was being carried out and women who had had such scans were not eligible (it was not clear how many women were excluded on this basis) and this compromised recruitment. Approximately half of the desired sample was recruited and the study did not have the required power to detect differences between groups.

RADIUS 1993

Methods	RCT after stratification by practice site.
Participants	15530 women. Inclusion criteria: women who did not have 'an indication for ultrasonography' based on medical disorder, uncertain gestational age, previous or current pregnancy complication, i.e. those who were eligible for inclusion were at low risk of adverse pregnancy outcome (and comprised 40% of the total population).
Interventions	Intervention group: (n = 7812) ultrasound screen at 18-20 and at 31-33 weeks' gestation. Control group: (n = 7718) selective ultrasonography for specific reasons only.

RADIUS 1993 *(Continued)*

	97% of women in the screened group had at least 2 ultrasound examinations; 55% of women in the control group had no scan at all. The mean number of scans was 2.2 (screened group) and 0.6 (controls). Ultrasound was to include a detailed study of fetal anatomy. Recruitment 1987-1991.
Outcomes	Fetal outcome and indices of care/intervention during pregnancy. The primary outcomes were fetal and neonatal mortality, and 'moderate or severe' neonatal morbidity.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	According to a "computer-based randomisation sequence" with stratification for practice site.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding (performance bias and detection bias) Women	High risk	Not attempted.
Blinding (performance bias and detection bias) Clinical staff	High risk	Clinical staff caring for the women were informed of the results of all ultrasound scans
Blinding (performance bias and detection bias) Outcome assessors	High risk	Ultrasound scan results were available in the case notes from which data were abstracted at the end of the trial
Incomplete outcome data (attrition bias) All outcomes	Low risk	Analysis for primary outcomes for all women not lost to follow-up (available case analysis). 1.6% were lost to follow-up and 0.8% had spontaneous miscarriages.
Selective reporting (reporting bias)	Unclear risk	None apparent.
Other bias	Unclear risk	Large numbers of women ineligible for inclusion (15,530 of 53,367 randomised).

Sweden 1988

Methods	RCT.
Participants	4997 women booking for antenatal care between October 1985 and March 1987 at 19 antenatal clinics in Stockholm, Sweden. Women were approximately 11 weeks' gestation at randomisation. Inclusion criteria: all consenting women at < 19 weeks who had not already had an ultrasound scan and who did not have one of a number of pre-specified indications for ultrasound (mainly uncertainties about gestational age, medical disorder, previous complications).
Interventions	Intervention group: planned routine ultrasound at about 15 weeks (range 13-19 weeks). Scans were carried out either in the ultrasound department or by trained midwives or obstetricians. BPD was measured and fetal viability and multiple pregnancy noted (98.7% had a scan as planned).

Sweden 1988 *(Continued)*

	Control groups: no routine scan unless indicated (4.1% had a scan before and 31% of control group women had an ultrasound scan after 19 weeks).
Outcomes	Major outcome; "neonatal morbidity" defined by admission to (and duration of stay in) neonatal ward. Follow-up data on neurodevelopmental outcome are available for around 70% of the sample at ages 8-9; these data were obtained by postal questionnaire. Data are also available on growth characteristics during childhood but not in a form that allows inclusion in the data tables; there was little difference between groups. School records of teenagers are available for 94% of singletons; data on school performance were not reported in a form that allowed us to include them in data tables.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Low risk	"opaque sealed envelope method."
Blinding (performance bias and detection bias) Women	High risk	Not attempted
Blinding (performance bias and detection bias) Clinical staff	High risk	Clinical staff looking after women were aware of the scan results
Blinding (performance bias and detection bias) Outcome assessors	High risk	Not mentioned.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	(32% of those otherwise eligible to participate were not included as they had indications for elective scanning.) 4997 women were randomised. There was small loss to follow-up for pregnancy outcomes but greater loss to follow-up and missing data for childhood developmental outcomes (> 50% attrition for some outcomes).
Selective reporting (reporting bias)	Unclear risk	Trial reported in multiple publications. We did not assess the trial protocol.
Other bias	Unclear risk	No baseline imbalance apparent. We have included all women randomised in the denominators for pregnancy outcomes.

Trondheim 1984

Methods	RCT.
Participants	1009 women attending their first antenatal visit at 25 general practitioners in Trondheim, Norway. Recruitment 1979-1980.
Interventions	Intervention group: ultrasound scans at 19 and 32 weeks' gestation. At the 19 week scan the BPD was measured to assess gestational age and to predict the EDD. Placental location and multiple pregnan-

Trondheim 1984 *(Continued)*

cies were noted. At the 32 week scan the mean abdominal diameter and the BPD were assessed, the placental location and presentation were noted.

Control group: no routine scans.

Outcomes	Reduction in post-term labour inductions. Birthweight and NICU admission, interventions in labour.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Described as "randomly selected" "randomised to two equally large groups".
Allocation concealment (selection bias)	Unclear risk	Described as "sealed envelope method".
Blinding (performance bias and detection bias) Women	High risk	Not attempted.
Blinding (performance bias and detection bias) Clinical staff	High risk	Not attempted
Blinding (performance bias and detection bias) Outcome assessors	Unclear risk	Partial blinding of outcome assessment for some outcomes "those who assessed neonatal outcomes did not know whether mothers were cases or controls".
Incomplete outcome data (attrition bias) All outcomes	Low risk	Very low levels of missing data.1009 women were randomised; of the 510 women in the intervention group there were 13 miscarriages (3 after the screening) and 1 woman was lost to follow-up; of 499 controls there were 19 miscarriages and 2 women were lost to follow-up.
Selective reporting (reporting bias)	Low risk	Not apparent.
Other bias	Low risk	None apparent.

Tygerberg 1996

Methods	RCT.
Participants	988 women attending clinics in Tygerberg, South Africa, where there was no policy for routine ultrasound and where many women did not have access to facilities for ultrasound. Recruitment between 1991-2. Randomisation at 18-24 weeks' with gestational age being estimated from menstrual history and clinical examination. Inclusion criteria: urban pregnant women attending for antenatal care before 24 weeks' gestation and planning to deliver in the Tygerberg area. Exclusion criteria: women aged over 37 and those that had already had an ultrasound. Women with increased risk of congenital abnormalities, with diabetes or rhesus sensitisation were also excluded.
Interventions	Intervention group: "routine ultrasound" by trained obstetric registrar. Results were recorded in notes.

Tygerberg 1996 *(Continued)*

	Control group: no routine ultrasound, selective ultrasound at the discretion of the managing clinician.
Outcomes	The study examined whether routine ultrasound would reduce the use of other antenatal services by improved dating and earlier diagnosis of outcomes and whether perinatal outcomes (gestational age at delivery, birthweight, perinatal mortality for babies over 28 weeks and NICU admissions) would be improved.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Low risk	"sealed, opaque envelopes."
Blinding (performance bias and detection bias) Women	High risk	Not attempted
Blinding (performance bias and detection bias) Clinical staff	High risk	Staff were aware of the results of the scan
Blinding (performance bias and detection bias) Outcome assessors	Unclear risk	Not mentioned.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	988 women randomised, 8% of women excluded after randomisation. Of 496 women randomised to the intervention group 3 were not pregnant and 36 were lost to follow-up. Of 492 controls 1 woman was not pregnant and 39 were lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	Not possible to assess study protocol.
Other bias	Low risk	No baseline imbalance apparent.

BPD: biparietal diameter
EDD: expected date of delivery
ITT: intention-to-treat
mins: minute
NICU: neonatal intensive care unit
RCT: randomised controlled trial

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

Study	Reason for exclusion
Belanger 1996	Study design: RCT. Intervention: scans at 16-20 weeks and 30-36 weeks versus scans only when clinically indicated in a low-risk population. Participants: 286 singleton infants.

	Outcomes: Bayley evaluations at 6 and 18 months of age giving 2 scores, Mental Development Index and Psychomotor Development.
	Did not report on any measures included in the review.
Bennett 2004	Study design: RCT. Intervention: first trimester estimation of CRL for gestational age estimation versus second trimester biometry estimation of gestational age. Participants: 218 women. Outcomes: rate of induction of labour for post-term pregnancy. This study was excluded as women were randomised to receive either a first or second trimester scan (scans in both arms of the trial were carried out before 24 weeks' gestation).
Duff 1993	In this study, both groups had scans in early pregnancy. RCT examining an early "dating" scan versus an early "dating" scan and a further scan at 34 weeks' gestation (1528 women randomised). Outcomes: fetal distress in labour, operative delivery, Apgar scores at delivery, birthweight, perinatal mortality and admission to NICU.
Georgsson Ohman 2010	This study looked at a sub-group of women who were part of a RCT assessing the timing of early scans in groups randomised to receive scans for fetal assessment at 12 to 14 versus 15 to 18 weeks' gestation. All women therefore received an early pregnancy scan.
Hong Kong	In this RCT, women in the intervention group received a detailed morphology scan at 12 to 14 weeks as part of scheduled nuchal scan. Women in the control group received nuchal scan at 12 to 14 weeks but no detailed morphology scan at this stage. Women in both groups received detailed morphology scans at 16-23 weeks. The study was excluded as women in both groups had an early scan.
Larsen 1992	Study design: randomised controlled study. Intervention: US estimation of fetal weight at 28 weeks and then every third week until delivery. Participants: 1000 women considered at risk of small-for-gestational-age fetus. Outcomes: number of interventions during pregnancy (elective delivery, admission to hospital), emergency intervention during labour and fetal outcome. All women had early US estimation of gestational age.
Leung 2006	Study design: randomised controlled study. Intervention: 2-dimensional US versus 2-dimensional US followed by 3-dimensional/4-dimensional US. Participants: 124 women at high risk of a fetal anomaly. Outcomes: maternal anxiety levels at first visit, 18 weeks' gestation (immediately after US examination) and 28 weeks' gestation.
Owen 1994	Study design: unclear. Intervention: monthly US estimation of fetal growth until 30 weeks and then fortnightly until delivery. Participants: 274 women with no risk factors for abnormal fetal growth. Outcomes: growth velocity. All women had early US estimation of gestational age.
Rustico 2005	Study design: RCT. Intervention: 2-dimensional versus 4-dimensional US in the second/third trimester of pregnancy. Participants: 100 women. Outcomes: ability of women to visualise fetal structures and movements.

Study	Reason for exclusion
Saltvedt 2006	In this study, the timing of early scans was examined in groups randomised to receive scans for fetal assessment at 12 to 14 versus 15 to 18 weeks' gestation.
Schifano 2010	Study design: RCT Intervention: Routine 2D scan in early pregnancy versus Routine 2D and 4D scan in early pregnancy Participants: 70 women attending a fetal medicine unit Outcomes: satisfaction with US scanning and maternal emotional status
Schwarzler 1999	This study examined the optimal timing of scans. Women were randomised to receive scans at 18, 20, or 22 weeks' gestation.
Votino 2012	In this study, all women received a first trimester US at 11 to 13 + 6 weeks. Women were randomised to 1 of 4 different groups using different methods to perform a detailed fetal cardiac examination.
Wald 1988	Study design: RCT. Intervention: routine scan in early pregnancy versus no scan. Participants: unclear. Outcomes: birthweight, type of labour, mode of delivery, NICU admission, gestation at delivery. Data not available after contact with authors.

CRL:crown-rump length
NICU: neonatal intensive care unit
RCT: randomised controlled trial
US: ultrasound

Characteristics of studies awaiting assessment [ordered by study ID]

Newcastle 2004

Methods	RCT.
Participants	15,209 women randomised.
Interventions	Intervention group: early US (11-14 weeks) for fetal abnormalities including anatomic assessment and nuchal translucency screening. Controls: 19 week scan.
Outcomes	Anxiety and depression costs. Anomalies diagnosed.
Notes	The results of the trial were published in a series of brief abstracts. Denominators for results were not provided.

RCT: randomised controlled trial
US: ultrasound

Characteristics of ongoing studies [ordered by study ID]

McClure 2014

Trial name or title	First look: a cluster-randomised trial of ultrasound to improve pregnancy outcomes in low-income settings.
Methods	Multi-country cluster-randomised trial.
Participants	Women attending for antenatal care in 58 clusters (approximately 500 births per cluster) in 5 low-income rural settings (Guatemala, Zambia, Kenya, Pakistan).
Interventions	US examination at 18-22 and at 32-36 weeks.
Outcomes	Primary outcome is a composite outcome of maternal mortality, maternal near miss, stillbirth and neonatal mortality. Secondary outcomes include rate of antenatal care utilisation and utilisation of health facilities for delivery by women with complicated pregnancies.
Starting date	April 2014
Contact information	Dr Elizabeth M McClure mcclure@rti.org
Notes	

US: ultrasound

DATA AND ANALYSES

Comparison 1. Routine/revealed versus selective/concealed ultrasound in early pregnancy

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Detection of fetal abnormality before 24 weeks' gestation	2	387	Risk Ratio (M-H, Fixed, 95% CI)	3.46 [1.67, 7.14]
2 Detection of multiple pregnancy by 24 to 26 weeks' gestation (number NOT detected)	7	295	Risk Ratio (M-H, Fixed, 95% CI)	0.07 [0.03, 0.17]
3 Induction of labour for 'post-term' pregnancy	8	25516	Risk Ratio (M-H, Random, 95% CI)	0.59 [0.42, 0.83]
4 Perinatal death (all babies)	10	35735	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.70, 1.12]
5 Perinatal death (excluding lethal malformations)	8	34331	Risk Ratio (M-H, Fixed, 95% CI)	0.96 [0.72, 1.27]
6 Detection of multiple pregnancy before labour (number NOT detected)	5	273	Risk Ratio (M-H, Fixed, 95% CI)	0.12 [0.03, 0.54]
7 Detection of major anomaly before birth	2	387	Risk Ratio (M-H, Fixed, 95% CI)	3.19 [1.99, 5.11]
8 Low birthweight (less than 2500 g)	8	19337	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.82, 1.33]
8.1 Singletons	4	15868	Risk Ratio (M-H, Random, 95% CI)	0.83 [0.60, 1.15]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
8.2 All babies (or not clear)	4	3469	Risk Ratio (M-H, Random, 95% CI)	1.28 [1.00, 1.64]
9 Very low birthweight (< 1500 g)	2	1584	Risk Ratio (M-H, Random, 95% CI)	1.26 [0.27, 5.82]
9.1 Singletons	0	0	Risk Ratio (M-H, Random, 95% CI)	0.0 [0.0, 0.0]
9.2 All babies or not clear	2	1584	Risk Ratio (M-H, Random, 95% CI)	1.26 [0.27, 5.82]
10 Small-for-gestational age	3	17105	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.81, 1.35]
11 Mean birthweight (g)	5	23213	Mean Difference (IV, Random, 95% CI)	10.67 [-19.77, 41.11]
12 Apgar score 7 or less at 5 minutes	4	3906	Risk Ratio (M-H, Random, 95% CI)	0.76 [0.33, 1.72]
13 Admission to neonatal intensive care unit (various definitions)	8	19088	Risk Ratio (M-H, Fixed, 95% CI)	0.95 [0.88, 1.02]
14 Impaired development (screened us-ing the Denver developmental screening test) at childhood follow-up	1	1657	Risk Ratio (M-H, Fixed, 95% CI)	0.95 [0.70, 1.30]
15 Poor oral reading at school	1	1993	Risk Ratio (M-H, Fixed, 95% CI)	1.02 [0.74, 1.41]
16 Poor reading comprehension at school	1	1984	Risk Ratio (M-H, Fixed, 95% CI)	0.82 [0.56, 1.22]
17 Poor spelling at school	1	1982	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.64, 1.13]
18 Poor arithmetic at school	1	1993	Risk Ratio (M-H, Fixed, 95% CI)	0.90 [0.61, 1.35]
19 Poor overall school performance	1	1993	Risk Ratio (M-H, Fixed, 95% CI)	0.96 [0.62, 1.47]
20 Dyslexia	1	603	Risk Ratio (M-H, Fixed, 95% CI)	0.77 [0.44, 1.34]
21 Reduced hearing in childhood	2	5418	Risk Ratio (M-H, Fixed, 95% CI)	0.90 [0.68, 1.20]
22 Reduced vision in childhood	2	5417	Risk Ratio (M-H, Fixed, 95% CI)	0.83 [0.68, 1.01]
23 Use of spectacles	2	5331	Risk Ratio (M-H, Fixed, 95% CI)	0.88 [0.74, 1.04]
24 Non right-handedness	2	4715	Risk Ratio (M-H, Random, 95% CI)	1.12 [0.92, 1.36]
25 Ambidexterity	1	1663	Risk Ratio (M-H, Fixed, 95% CI)	1.23 [0.92, 1.63]
26 Appropriately timed serum screening tests (number having repeat screening)	1	602	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.45, 1.76]
27 Appropriately timed anomaly scan (18 to 22 weeks)(number NOT appropri-ately timed)	1	602	Risk Ratio (M-H, Fixed, 95% CI)	0.77 [0.55, 1.08]
28 Termination of pregnancy for fetal abnormality	5	28256	Risk Ratio (M-H, Fixed, 95% CI)	2.23 [1.10, 4.54]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
29 Number of antenatal visits	2	9502	Mean Difference (IV, Random, 95% CI)	0.16 [-0.33, 0.65]
30 Antenatal hospital admission	6	17785	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.92, 1.18]
31 Induction of labour for any reason	7	24790	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.63, 0.97]
32 Caesarean section	5	22193	Risk Ratio (M-H, Fixed, 95% CI)	1.05 [0.98, 1.12]
33 Mother not satisfied with care (worried about pregnancy)	1	634	Risk Ratio (M-H, Fixed, 95% CI)	0.80 [0.65, 0.99]
34 Subgroup analysis by timing of scan: detection of multiple pregnancy by 24-26 weeks' gestation (number not detected)	7	295	Risk Ratio (M-H, Fixed, 95% CI)	0.07 [0.03, 0.17]
34.1 Ultrasound planned before 14 weeks	1	9	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.05, 16.36]
34.2 Ultrasound after 14 weeks	6	286	Risk Ratio (M-H, Fixed, 95% CI)	0.06 [0.02, 0.16]
35 Subgroup analysis: induction of labour for 'post-term' pregnancy (early and later scans)	7	24712	Risk Ratio (M-H, Random, 95% CI)	0.61 [0.42, 0.88]
35.1 Scan before 14 weeks	2	1278	Risk Ratio (M-H, Random, 95% CI)	0.99 [0.67, 1.46]
35.2 Scan after 14 weeks	5	23434	Risk Ratio (M-H, Random, 95% CI)	0.49 [0.31, 0.77]
36 Subgroup analysis: perinatal death (earlier and late scans)	9	34923	Risk Ratio (M-H, Fixed, 95% CI)	0.87 [0.68, 1.11]
36.1 Scan before 14 weeks	2	1416	Risk Ratio (M-H, Fixed, 95% CI)	0.73 [0.23, 2.30]
36.2 Scan after 14 weeks	7	33507	Risk Ratio (M-H, Fixed, 95% CI)	0.87 [0.68, 1.13]
37 Subgroup analysis: detection of multiple pregnancy before 24 weeks (number not detected; concealed results	7	295	Risk Ratio (M-H, Fixed, 95% CI)	0.07 [0.03, 0.17]
37.1 Concealed results for controls	1	11	Risk Ratio (M-H, Fixed, 95% CI)	0.17 [0.01, 2.92]
37.2 Selective scan for controls	6	284	Risk Ratio (M-H, Fixed, 95% CI)	0.07 [0.02, 0.17]
38 Subgroup analysis: perinatal death. Concealed results	9	34923	Risk Ratio (M-H, Fixed, 95% CI)	0.87 [0.68, 1.11]
38.1 Concealed results for controls	1	1073	Risk Ratio (M-H, Fixed, 95% CI)	1.66 [0.40, 6.93]
38.2 Selective scans for controls	8	33850	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.66, 1.09]

Analysis 1.1. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 1 Detection of fetal abnormality before 24 weeks' gestation.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Alesund 1984	2/19	0/18		5.66%	4.75[0.24,92.65]
RADIUS 1993	31/187	8/163		94.34%	3.38[1.6,7.14]
Total (95% CI)	**206**	**181**		**100%**	**3.46[1.67,7.14]**
Total events: 33 (Routine scan), 8 (Control)					
Heterogeneity: Tau²=0; Chi²=0.05, df=1(P=0.83); I²=0%					
Test for overall effect: Z=3.35(P=0)					

Favours control 0.01 0.1 1 10 100 Favours routine scan

Analysis 1.2. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 2 Detection of multiple pregnancy by 24 to 26 weeks' gestation (number NOT detected).

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Alesund 1984	0/10	7/7		14.05%	0.05[0,0.73]
Helsinki 1990	0/36	9/38		14.96%	0.06[0,0.92]
London 1982	0/6	2/5		4.35%	0.17[0.01,2.92]
Missouri 1990	0/2	1/7		1.32%	0.89[0.05,16.36]
RADIUS 1993	1/68	23/61		39.22%	0.04[0.01,0.28]
Sweden 1988	0/25	13/20		24.16%	0.03[0,0.47]
Trondheim 1984	1/6	1/4		1.94%	0.67[0.06,7.85]
Total (95% CI)	**153**	**142**		**100%**	**0.07[0.03,0.17]**
Total events: 2 (Routine scan), 56 (Control)					
Heterogeneity: Tau²=0; Chi²=7.32, df=6(P=0.29); I²=18.09%					
Test for overall effect: Z=5.68(P<0.0001)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.3. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 3 Induction of labour for 'post-term' pregnancy.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Alesund 1984	13/774	44/750		12.87%	0.29[0.16,0.53]
Johannesburg 2007	6/416	14/388		8.21%	0.4[0.16,1.03]
Missouri 1990	28/402	31/413		15%	0.93[0.57,1.52]
Oxford 2006	19/233	17/230		12.57%	1.1[0.59,2.07]
RADIUS 1993	123/7617	161/7534		19.68%	0.76[0.6,0.95]
Sweden 1988	41/2389	88/2412		17.37%	0.47[0.33,0.68]
Trondheim 1984	14/496	19/478		11.75%	0.71[0.36,1.4]
Tygerberg 1996	1/493	14/491		2.55%	0.07[0.01,0.54]
Total (95% CI)	**12820**	**12696**		**100%**	**0.59[0.42,0.83]**
Total events: 245 (Routine scan), 388 (Control)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Random, 95% CI		M-H, Random, 95% CI
Heterogeneity: Tau²=0.14; Chi²=22.16, df=7(P=0); I²=68.42%					
Test for overall effect: Z=3.01(P=0)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.4. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 4 Perinatal death (all babies).

Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Fixed, 95% CI		M-H, Fixed, 95% CI
Adelaide 1999	3/296	3/298		2.03%	1.01[0.2,4.95]
Alesund 1984	6/794	10/765		6.91%	0.58[0.21,1.58]
Helsinki 1990	20/4422	39/4350		26.68%	0.5[0.29,0.86]
Johannesburg 2007	18/422	16/390		11.28%	1.04[0.54,2.01]
London 1982	5/537	3/536		2.04%	1.66[0.4,6.93]
Missouri 1990	2/402	4/420		2.65%	0.52[0.1,2.84]
RADIUS 1993	52/7685	41/7596		27.98%	1.25[0.83,1.89]
Sweden 1988	12/2415	12/2432		8.11%	1.01[0.45,2.24]
Trondheim 1984	5/503	5/482		3.47%	0.96[0.28,3.29]
Tygerberg 1996	9/496	13/494		8.84%	0.69[0.3,1.6]
Total (95% CI)	**17972**	**17763**		**100%**	**0.89[0.7,1.12]**
Total events: 132 (Routine scan), 146 (Control)					
Heterogeneity: Tau²=0; Chi²=9.52, df=9(P=0.39); I²=5.44%					
Test for overall effect: Z=1.01(P=0.31)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.5. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 5 Perinatal death (excluding lethal malformations).

Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Fixed, 95% CI		M-H, Fixed, 95% CI
Alesund 1984	6/794	7/765		7.46%	0.83[0.28,2.45]
Helsinki 1990	17/4422	28/4350		29.55%	0.6[0.33,1.09]
London 1982	4/537	3/536		3.14%	1.33[0.3,5.92]
Missouri 1990	2/404	4/420		4.11%	0.52[0.1,2.82]
RADIUS 1993	42/7685	27/7596		28.43%	1.54[0.95,2.49]
Sweden 1988	10/2415	10/2432		10.43%	1.01[0.42,2.42]
Trondheim 1984	5/503	5/482		5.35%	0.96[0.28,3.29]
Tygerberg 1996	6/496	11/494		11.54%	0.54[0.2,1.46]
Total (95% CI)	**17256**	**17075**		**100%**	**0.96[0.72,1.27]**
Total events: 92 (Routine scan), 95 (Control)					
Heterogeneity: Tau²=0; Chi²=8.11, df=7(P=0.32); I²=13.68%					
Test for overall effect: Z=0.3(P=0.76)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.6. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 6 Detection of multiple pregnancy before labour (number NOT detected).

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Alesund 1984	0/10	3/7		29.31%	0.1[0.01,1.74]
Helsinki 1990	0/36	0/38			Not estimable
Missouri 1990	0/2	1/7		5.92%	0.89[0.05,16.36]
RADIUS 1993	0/68	8/61		64.77%	0.05[0,0.9]
Sweden 1988	0/24	0/20			Not estimable
Total (95% CI)	140	133		100%	0.12[0.03,0.54]

Total events: 0 (Routine scan), 12 (Control)

Heterogeneity: Tau²=0; Chi²=2.17, df=2(P=0.34); I²=7.79%

Test for overall effect: Z=2.74(P=0.01)

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.7. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 7 Detection of major anomaly before birth.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Alesund 1984	2/19	0/18		2.6%	4.75[0.24,92.65]
RADIUS 1993	65/187	18/163		97.4%	3.15[1.95,5.08]
Total (95% CI)	206	181		100%	3.19[1.99,5.11]

Total events: 67 (Routine scan), 18 (Control)

Heterogeneity: Tau²=0; Chi²=0.07, df=1(P=0.79); I²=0%

Test for overall effect: Z=4.82(P<0.0001)

Favours control 0.01 0.1 1 10 100 Favours routine scan

Analysis 1.8. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 8 Low birthweight (less than 2500 g).

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
1.8.1 Singletons					
Alesund 1984	27/774	26/750		10.23%	1.01[0.59,1.71]
Helsinki 1990	130/4350	120/4274		16.88%	1.06[0.83,1.36]
Sweden 1988	59/2364	95/2392		14.99%	0.63[0.46,0.87]
Trondheim 1984	11/490	17/474		6.82%	0.63[0.3,1.32]
Subtotal (95% CI)	7978	7890		48.92%	0.83[0.6,1.15]

Total events: 227 (Routine scan), 258 (Control)

Heterogeneity: Tau²=0.06; Chi²=7.62, df=3(P=0.05); I²=60.64%

Test for overall effect: Z=1.1(P=0.27)

1.8.2 All babies (or not clear)					
Adelaide 1999	22/296	11/298		7.36%	2.01[0.99,4.08]
Johannesburg 2007	61/422	54/390		14.51%	1.04[0.74,1.47]

Favours routine scan 0.1 0.2 0.5 1 2 5 10 Favours control

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
London 1982	48/537	44/536		13.23%	1.09[0.74,1.61]
Tygerberg 1996	102/496	68/494		15.99%	1.49[1.13,1.98]
Subtotal (95% CI)	**1751**	**1718**		**51.08%**	**1.28[1,1.64]**
Total events: 233 (Routine scan), 177 (Control)					
Heterogeneity: Tau²=0.02; Chi²=4.79, df=3(P=0.19); I²=37.41%					
Test for overall effect: Z=2(P=0.05)					
Total (95% CI)	**9729**	**9608**		**100%**	**1.04[0.82,1.33]**
Total events: 460 (Routine scan), 435 (Control)					
Heterogeneity: Tau²=0.07; Chi²=21.18, df=7(P=0); I²=66.95%					
Test for overall effect: Z=0.35(P=0.73)					
Test for subgroup differences: Chi²=4.34, df=1 (P=0.04), I²=76.95%					

Favours routine scan 0.1 0.2 0.5 1 2 5 10 Favours control

Analysis 1.9. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 9 Very low birthweight (< 1500 g).

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
1.9.1 Singletons					
Subtotal (95% CI)	**0**	**0**			**Not estimable**
Total events: 0 (Routine scan), 0 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Not applicable					
1.9.2 All babies or not clear					
Adelaide 1999	4/296	1/298		30.83%	4.03[0.45,35.82]
Tygerberg 1996	12/496	16/494		69.17%	0.75[0.36,1.56]
Subtotal (95% CI)	**792**	**792**		**100%**	**1.26[0.27,5.82]**
Total events: 16 (Routine scan), 17 (Control)					
Heterogeneity: Tau²=0.74; Chi²=2.07, df=1(P=0.15); I²=51.79%					
Test for overall effect: Z=0.29(P=0.77)					
Total (95% CI)	**792**	**792**		**100%**	**1.26[0.27,5.82]**
Total events: 16 (Routine scan), 17 (Control)					
Heterogeneity: Tau²=0.74; Chi²=2.07, df=1(P=0.15); I²=51.79%					
Test for overall effect: Z=0.29(P=0.77)					
Test for subgroup differences: Not applicable					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.10. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 10 Small-for-gestational age.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
RADIUS 1993	176/7617	170/7534		51.99%	1.02[0.83,1.26]
Trondheim 1984	41/490	27/474		21.81%	1.47[0.92,2.35]

Favours routine scan 0.01 0.1 1 10 100 Favours control

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Tygerberg 1996	38/496	46/494		26.2%	0.82[0.55,1.24]
Total (95% CI)	**8603**	**8502**		**100%**	**1.05[0.81,1.35]**

Total events: 255 (Routine scan), 243 (Control)

Heterogeneity: Tau²=0.02; Chi²=3.35, df=2(P=0.19); I²=40.25%

Test for overall effect: Z=0.34(P=0.73)

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.11. Comparison 1 Routine/revealed versus selective/ concealed ultrasound in early pregnancy, Outcome 11 Mean birthweight (g).

Study or subgroup	Routine scan N	Mean(SD)	Control N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Alesund 1984	774	3602 (562.1)	750	3563 (562.1)		16.78%	39[-17.45,95.45]
Johannesburg 2007	422	2944 (618)	390	2962 (631)		9.52%	-18[-104.01,68.01]
RADIUS 1993	7617	3433 (544)	7514	3429 (535)		35.21%	4[-13.19,21.19]
Sweden 1988	2364	3521 (527)	2392	3479 (557)		28.2%	42[11.19,72.81]
Tygerberg 1996	496	2890 (627.3)	494	2962 (682.2)		10.29%	-72[-153.65,9.65]
Total ***	**11673**		**11540**			**100%**	**10.67[-19.77,41.11]**

Heterogeneity: Tau²=608.23; Chi²=9.87, df=4(P=0.04); I²=59.49%

Test for overall effect: Z=0.69(P=0.49)

Favours control -100 -50 0 50 100 Favours routine scan

Analysis 1.12. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 12 Apgar score 7 or less at 5 minutes.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Adelaide 1999	4/296	5/298		23.29%	0.81[0.22,2.97]
Alesund 1984	3/774	10/750		23.69%	0.29[0.08,1.05]
Missouri 1990	2/404	4/420		16.59%	0.52[0.1,2.82]
Trondheim 1984	15/490	9/474		36.42%	1.61[0.71,3.65]
Total (95% CI)	**1964**	**1942**		**100%**	**0.76[0.33,1.72]**

Total events: 24 (Routine scan), 28 (Control)

Heterogeneity: Tau²=0.3; Chi²=5.37, df=3(P=0.15); I²=44.17%

Test for overall effect: Z=0.67(P=0.51)

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.13. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 13 Admission to neonatal intensive care unit (various definitions).

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Adelaide 1999	187/296	187/298		17.84%	1.01[0.89,1.14]
Alesund 1984	63/774	65/750		6.32%	0.94[0.67,1.31]
Helsinki 1990	366/4350	371/4274		35.82%	0.97[0.84,1.11]
Johannesburg 2007	30/422	25/390		2.49%	1.11[0.66,1.85]
Missouri 1990	25/404	31/420		2.91%	0.84[0.5,1.39]
Sweden 1988	231/2364	275/2392		26.16%	0.85[0.72,1]
Trondheim 1984	17/490	22/474		2.14%	0.75[0.4,1.39]
Tygerberg 1996	73/496	66/494		6.33%	1.1[0.81,1.5]
Total (95% CI)	**9596**	**9492**		**100%**	**0.95[0.88,1.02]**

Total events: 992 (Routine scan), 1042 (Control)

Heterogeneity: Tau²=0; Chi²=4.78, df=7(P=0.69); I²=0%

Test for overall effect: Z=1.39(P=0.17)

Favours routine scan 0.1 0.2 0.5 1 2 5 10 Favours control

Analysis 1.14. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 14 Impaired development (screened using the Denver developmental screening test) at childhood follow-up.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Norway 1993	75/859	73/798		100%	0.95[0.7,1.3]
Total (95% CI)	**859**	**798**		**100%**	**0.95[0.7,1.3]**

Total events: 75 (Routine scan), 73 (Control)

Heterogeneity: Tau²=0; Chi²=0, df=0(P<0.0001); I²=100%

Test for overall effect: Z=0.3(P=0.77)

Favours experimental 0.01 0.1 1 10 100 Favours control

Analysis 1.15. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 15 Poor oral reading at school.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Norway 1993	70/1007	67/986		100%	1.02[0.74,1.41]
Total (95% CI)	**1007**	**986**		**100%**	**1.02[0.74,1.41]**

Total events: 70 (Routine scan), 67 (Control)

Heterogeneity: Not applicable

Test for overall effect: Z=0.14(P=0.89)

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.16. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 16 Poor reading comprehension at school.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Norway 1993	44/1005	52/979		100%	0.82[0.56,1.22]
Total (95% CI)	**1005**	**979**		**100%**	**0.82[0.56,1.22]**
Total events: 44 (Routine scan), 52 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.97(P=0.33)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.17. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 17 Poor spelling at school.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Norway 1993	80/1001	92/981		100%	0.85[0.64,1.13]
Total (95% CI)	**1001**	**981**		**100%**	**0.85[0.64,1.13]**
Total events: 80 (Routine scan), 92 (Control)					
Heterogeneity: Tau²=0; Chi²=0, df=0(P<0.0001); I²=100%					
Test for overall effect: Z=1.09(P=0.27)					

Favours routine care 0.01 0.1 1 10 100 Favours control

Analysis 1.18. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 18 Poor arithmetic at school.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Norway 1993	44/1004	48/989		100%	0.9[0.61,1.35]
Total (95% CI)	**1004**	**989**		**100%**	**0.9[0.61,1.35]**
Total events: 44 (Routine scan), 48 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.5(P=0.62)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.19. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 19 Poor overall school performance.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Norway 1993	40/1006	41/987		100%	0.96[0.62,1.47]
Total (95% CI)	**1006**	**987**		**100%**	**0.96[0.62,1.47]**

Favours routine scan 0.01 0.1 1 10 100 Favours control

Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Fixed, 95% CI		M-H, Fixed, 95% CI
Total events: 40 (Routine scan), 41 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.2(P=0.84)					

Favours routine scan   0.01   0.1   1   10   100   Favours control

Analysis 1.20. Comparison 1 Routine/revealed versus selective/ concealed ultrasound in early pregnancy, Outcome 20 Dyslexia.

Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Fixed, 95% CI		M-H, Fixed, 95% CI
Norway 1993	21/309	26/294		100%	0.77[0.44,1.34]
Total (95% CI)	**309**	**294**		**100%**	**0.77[0.44,1.34]**
Total events: 21 (Routine scan), 26 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.93(P=0.35)					

Favours routine scan   0.01   0.1   1   10   100   Favours control

Analysis 1.21. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 21 Reduced hearing in childhood.

Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Fixed, 95% CI		M-H, Fixed, 95% CI
Norway 1993	32/1112	38/1041		40.94%	0.79[0.5,1.25]
Sweden 1988	56/1651	56/1614		59.06%	0.98[0.68,1.41]
Total (95% CI)	**2763**	**2655**		**100%**	**0.9[0.68,1.2]**
Total events: 88 (Routine scan), 94 (Control)					
Heterogeneity: Tau²=0; Chi²=0.51, df=1(P=0.47); I²=0%					
Test for overall effect: Z=0.72(P=0.47)					

Favours routine scan   0.01   0.1   1   10   100   Favours control

Analysis 1.22. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 22 Reduced vision in childhood.

Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Fixed, 95% CI		M-H, Fixed, 95% CI
Norway 1993	68/1113	72/1039		36.89%	0.88[0.64,1.21]
Sweden 1988	103/1651	126/1614		63.11%	0.8[0.62,1.03]
Total (95% CI)	**2764**	**2653**		**100%**	**0.83[0.68,1.01]**
Total events: 171 (Routine scan), 198 (Control)					
Heterogeneity: Tau²=0; Chi²=0.22, df=1(P=0.64); I²=0%					
Test for overall effect: Z=1.85(P=0.06)					

Favours routine scan   0.01   0.1   1   10   100   Favours control

Analysis 1.23. Comparison 1 Routine/revealed versus selective/ concealed ultrasound in early pregnancy, Outcome 23 Use of spectacles.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Norway 1993	82/1069	95/997		39.48%	0.81[0.61,1.07]
Sweden 1988	141/1651	149/1614		60.52%	0.93[0.74,1.15]
Total (95% CI)	**2720**	**2611**		**100%**	**0.88[0.74,1.04]**
Total events: 223 (Routine scan), 244 (Control)					
Heterogeneity: Tau²=0; Chi²=0.58, df=1(P=0.45); I²=0%					
Test for overall effect: Z=1.47(P=0.14)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.24. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 24 Non right-handedness.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Norway 1993	162/861	120/802		43.39%	1.26[1.01,1.56]
Sweden 1988	253/1544	240/1508		56.61%	1.03[0.88,1.21]
Total (95% CI)	**2405**	**2310**		**100%**	**1.12[0.92,1.36]**
Total events: 415 (Routine scan), 360 (Control)					
Heterogeneity: Tau²=0.01; Chi²=2.11, df=1(P=0.15); I²=52.72%					
Test for overall effect: Z=1.17(P=0.24)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.25. Comparison 1 Routine/revealed versus selective/ concealed ultrasound in early pregnancy, Outcome 25 Ambidexterity.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Norway 1993	100/861	76/802		100%	1.23[0.92,1.63]
Total (95% CI)	**861**	**802**		**100%**	**1.23[0.92,1.63]**
Total events: 100 (Routine scan), 76 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=1.41(P=0.16)					

Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.26. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 26 Appropriately timed serum screening tests (number having repeat screening).

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Adelaide 1999	15/299	17/303		100%	0.89[0.45,1.76]
Total (95% CI)	**299**	**303**		**100%**	**0.89[0.45,1.76]**

Total events: 15 (Routine scan), 17 (Control)

Heterogeneity: Not applicable

Test for overall effect: Z=0.32(P=0.75)

Favours scan 0.01 0.1 1 10 100 Favours control

Analysis 1.27. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 27 Appropriately timed anomaly scan (18 to 22 weeks)(number NOT appropriately timed).

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Adelaide 1999	48/299	63/303		100%	0.77[0.55,1.08]
Total (95% CI)	**299**	**303**		**100%**	**0.77[0.55,1.08]**

Total events: 48 (Routine scan), 63 (Control)

Heterogeneity: Not applicable

Test for overall effect: Z=1.49(P=0.14)

Favours scan 0.01 0.1 1 10 100 Favours control

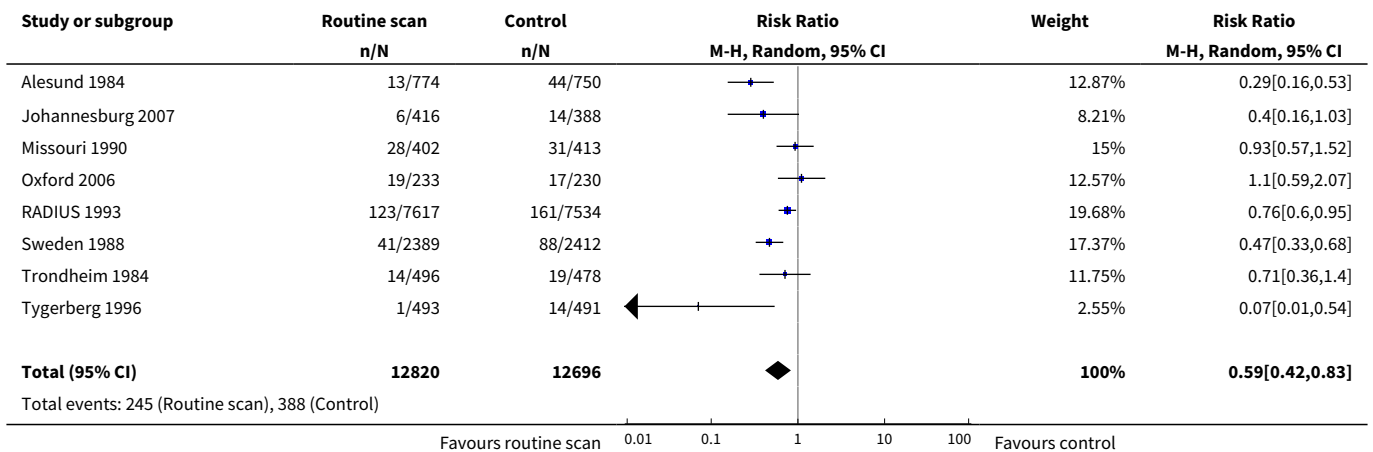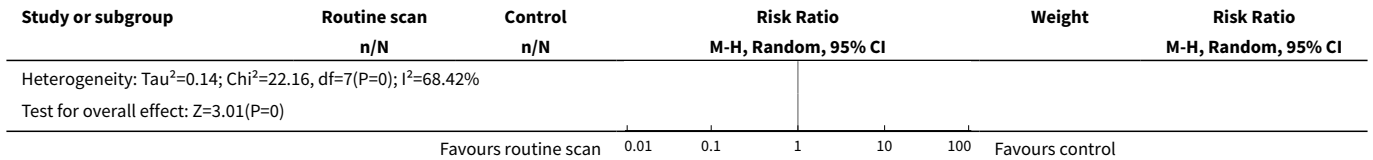### Analysis 1.28. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 28 Termination of pregnancy for fetal abnormality.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Alesund 1984	2/825	0/803		4.57%	4.87[0.23,101.22]
Helsinki 1990	11/4691	4/4619		36.36%	2.71[0.86,8.5]
Johannesburg 2007	1/416	0/388		4.67%	2.8[0.11,68.49]
RADIUS 1993	9/7812	5/7718		45.37%	1.78[0.6,5.3]
Tygerberg 1996	1/493	1/491		9.04%	1[0.06,15.88]
Total (95% CI)	**14237**	**14019**		**100%**	**2.23[1.1,4.54]**

Total events: 24 (Routine scan), 10 (Control)

Heterogeneity: Tau²=0; Chi²=0.88, df=4(P=0.93); I²=0%

Test for overall effect: Z=2.22(P=0.03)

More events control 0.01 0.1 1 10 100 More events routine scan

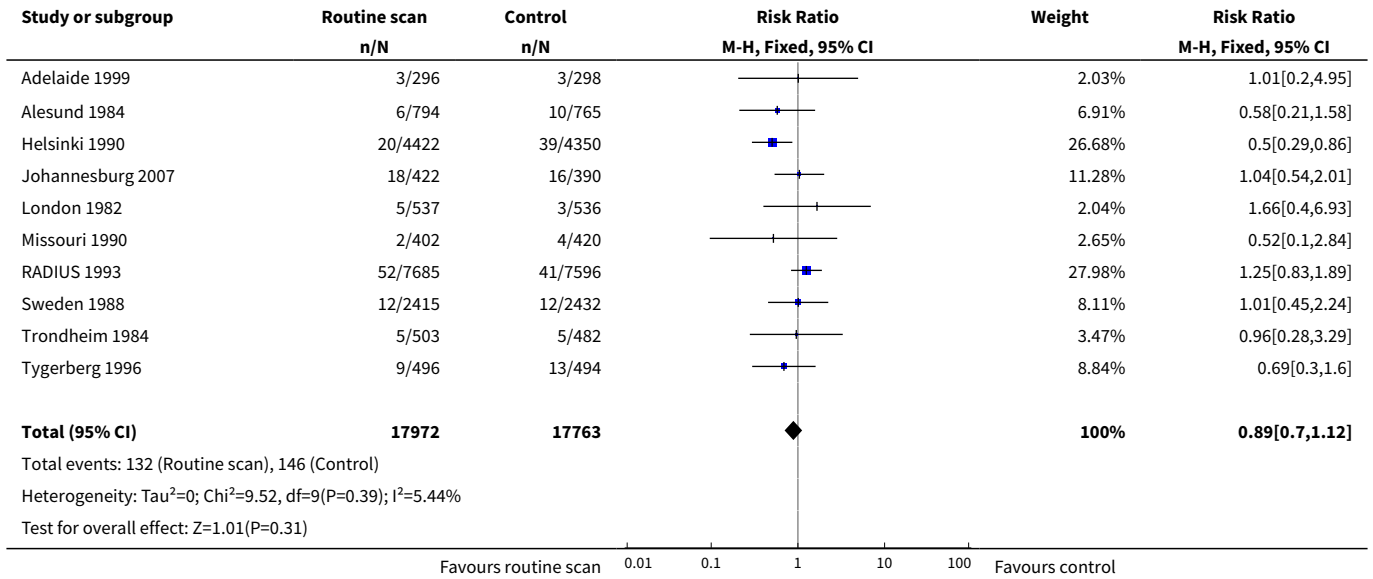### Analysis 1.29. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 29 Number of antenatal visits.

Study or subgroup	Routine scan N	Routine scan Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Helsinki 1990	4386	12.9 (3.3)	4312	12.5 (3.2)		52.65%	0.4[0.26,0.54]
Johannesburg 2007	416	5.8 (1.8)	388	5.9 (2)		47.35%	-0.1[-0.36,0.16]
Total ***	**4802**		**4700**			**100%**	**0.16[-0.33,0.65]**

Heterogeneity: Tau²=0.11; Chi²=10.89, df=1(P=0); I²=90.82%

Test for overall effect: Z=0.65(P=0.51)

Favours routine care -100 -50 0 50 100 Favours control

Analysis 1.30. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 30 Antenatal hospital admission.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Alesund 1984	148/774	178/750		17.39%	0.81[0.66,0.98]
Helsinki 1990	1489/4386	1469/4312		27.68%	1[0.94,1.06]
Johannesburg 2007	68/416	51/388		9.54%	1.24[0.89,1.74]
Sweden 1988	338/2389	335/2412		21.59%	1.02[0.89,1.17]
Trondheim 1984	79/496	46/478		9.33%	1.66[1.18,2.33]
Tygerberg 1996	110/493	105/491		14.47%	1.04[0.82,1.32]
Total (95% CI)	**8954**	**8831**		**100%**	**1.04[0.92,1.18]**

Total events: 2232 (Routine scan), 2184 (Control)

Heterogeneity: Tau²=0.01; Chi²=15, df=5(P=0.01); I²=66.67%

Test for overall effect: Z=0.61(P=0.54)

Favours routine scan 0.1 0.2 0.5 1 2 5 10 Favours control

Analysis 1.31. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 31 Induction of labour for any reason.

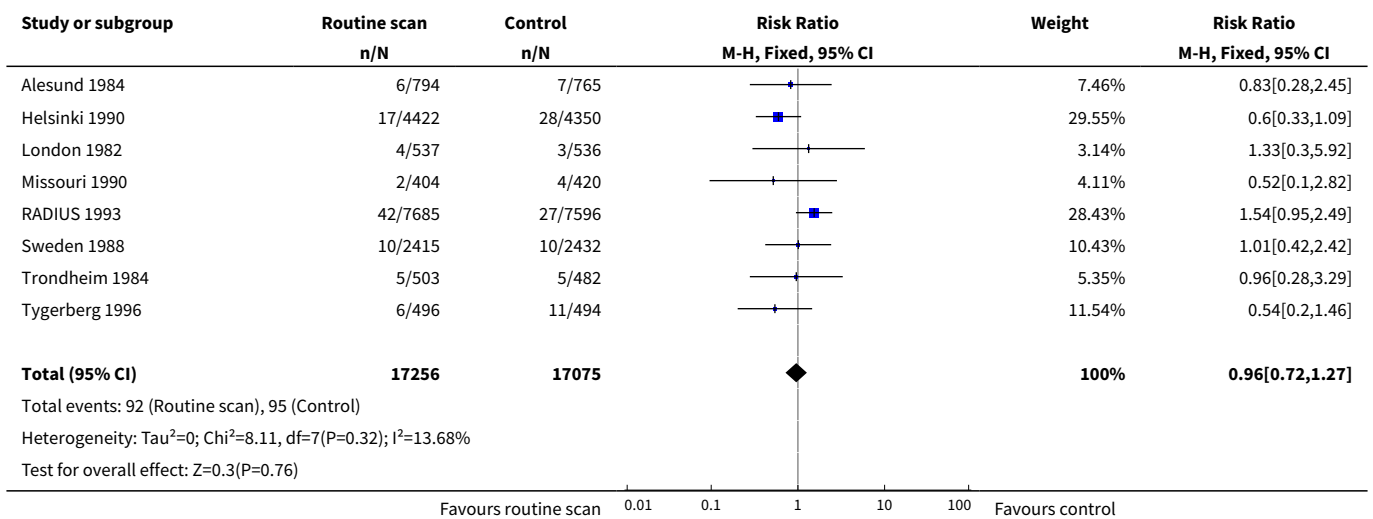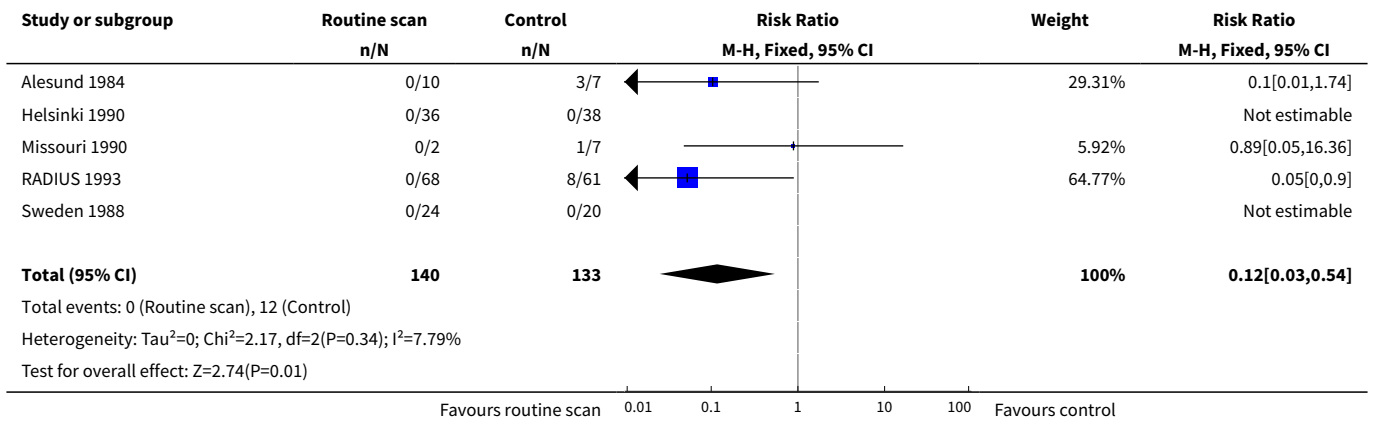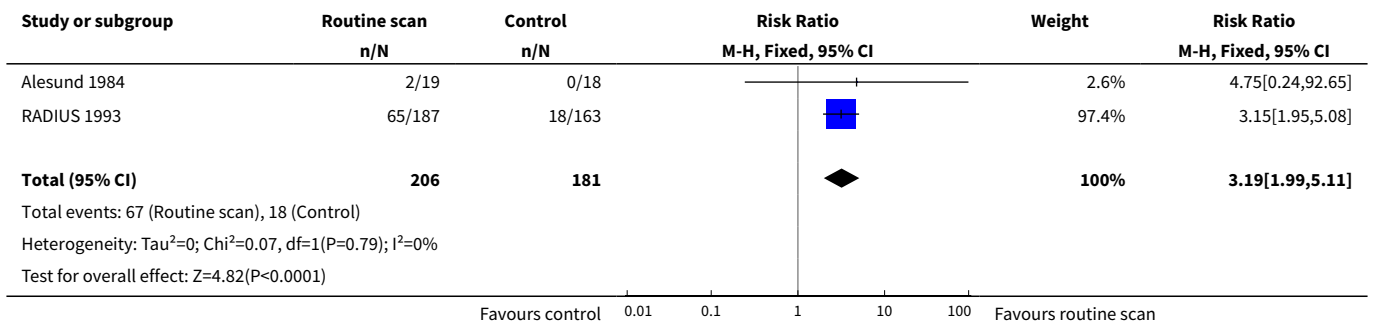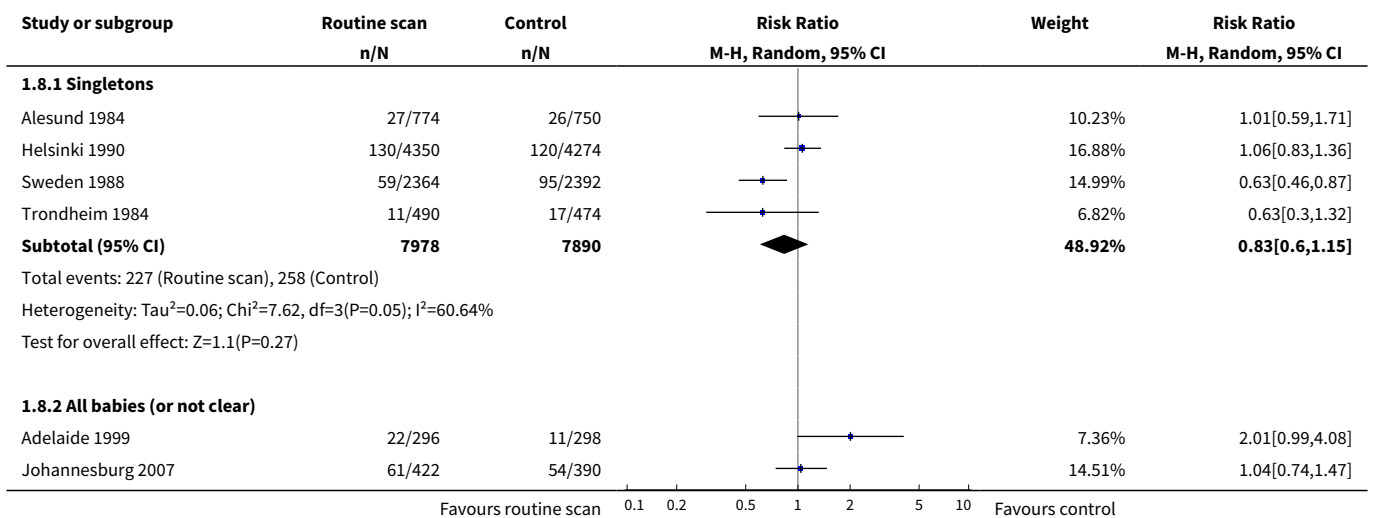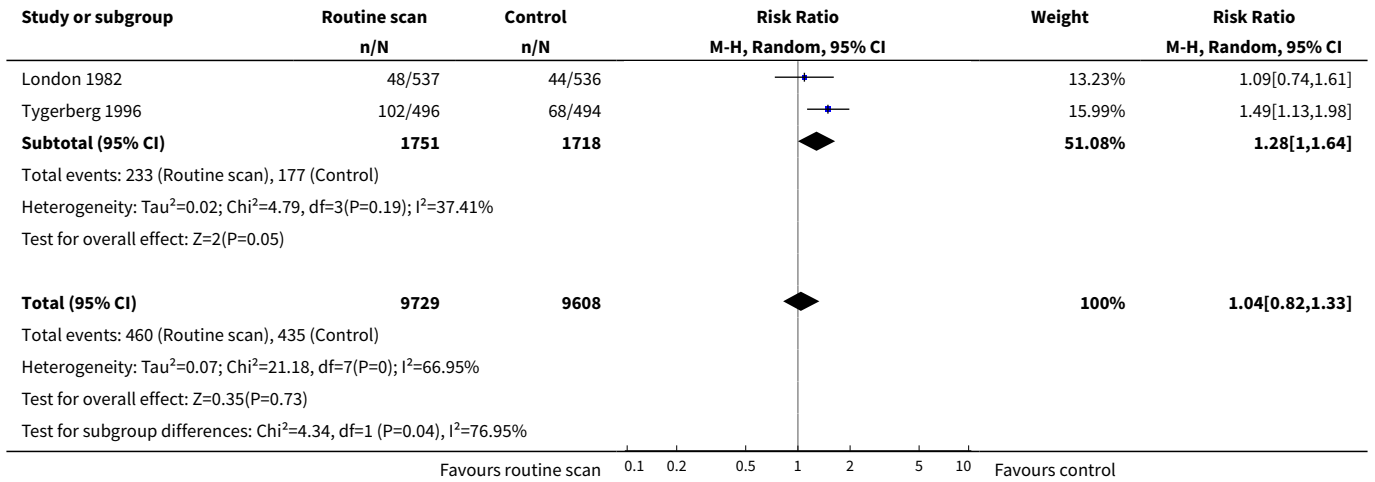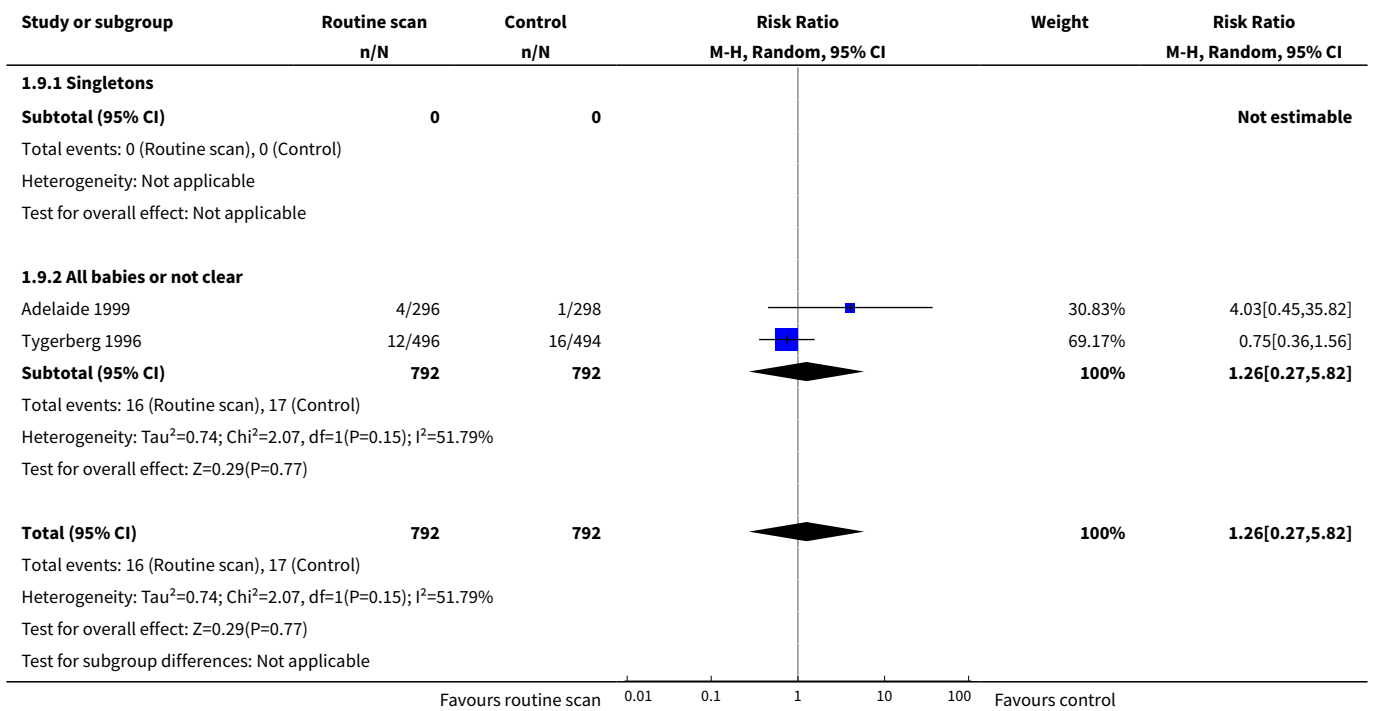| Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Alesund 1984	34/774	77/750		11.76%	0.43[0.29,0.63]
London 1982	104/531	107/531		15.51%	0.97[0.76,1.24]
Missouri 1990	91/402	103/413		15.4%	0.91[0.71,1.16]
Oxford 2006	34/233	46/230		11.45%	0.73[0.49,1.09]
RADIUS 1993	1912/7617	1858/7534		19.08%	1.02[0.96,1.08]
Sweden 1988	140/2389	218/2412		16.43%	0.65[0.53,0.8]
Trondheim 1984	32/496	38/478		10.36%	0.81[0.52,1.28]
Total (95% CI)	**12442**	**12348**		**100%**	**0.78[0.63,0.97]**

Total events: 2347 (Routine scan), 2447 (Control)

Heterogeneity: Tau²=0.06; Chi²=37.57, df=6(P<0.0001); I²=84.03%

Test for overall effect: Z=2.24(P=0.03)
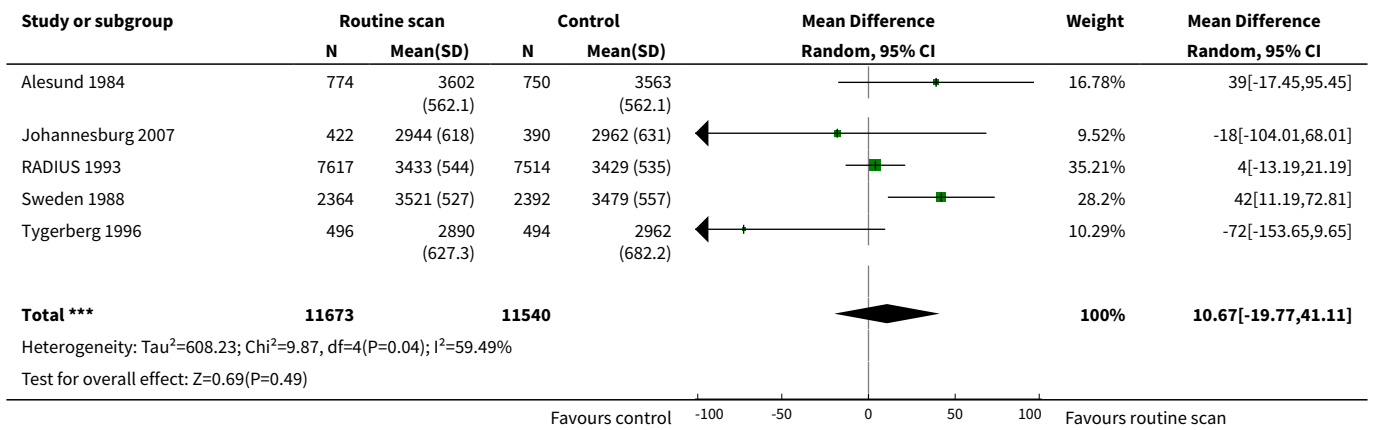
Favours routine scan 0.01 0.1 1 10 100 Favours control

Analysis 1.32. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 32 Caesarean section.
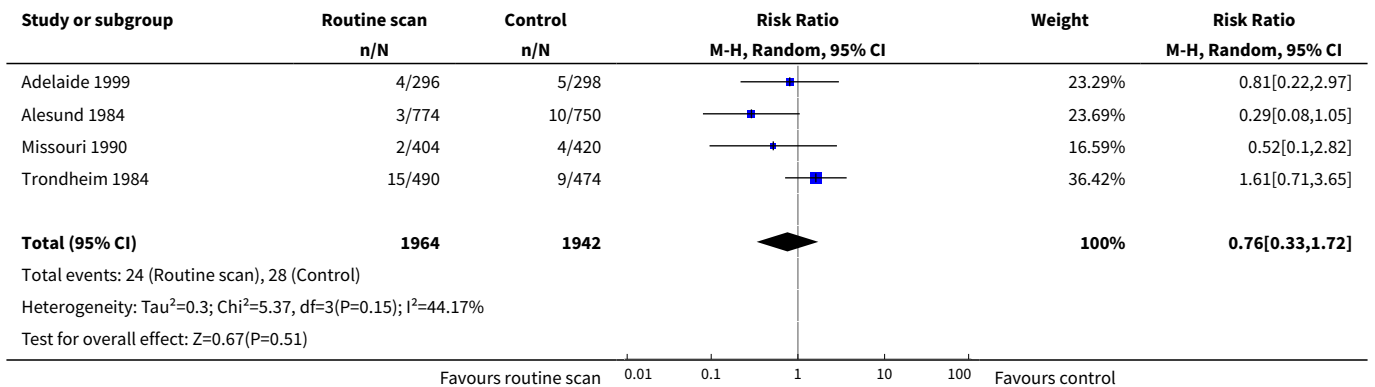
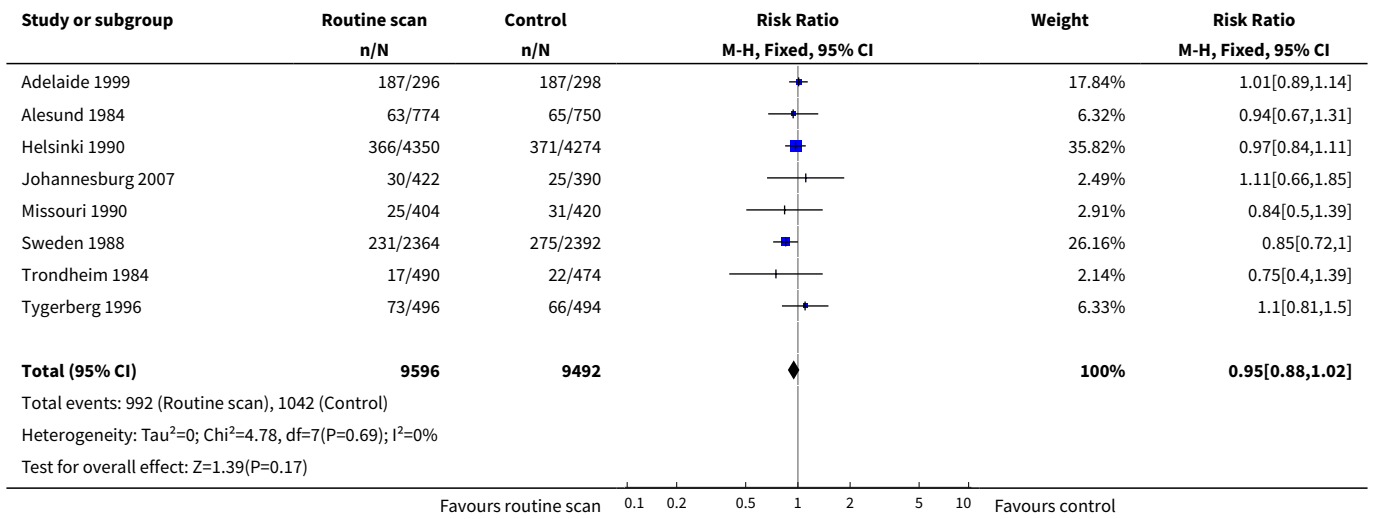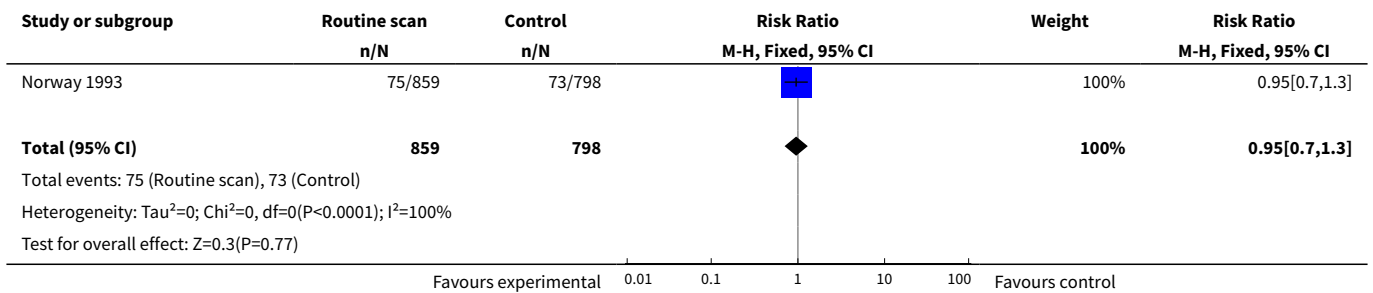| Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Johannesburg 2007	60/416	59/388		4.17%	0.95[0.68,1.32]
Oxford 2006	33/233	27/230		1.86%	1.21[0.75,1.94]
RADIUS 1993	1205/7617	1135/7534		78.02%	1.05[0.97,1.13]
Sweden 1988	210/2389	217/2412		14.76%	0.98[0.82,1.17]
Trondheim 1984	29/496	17/478		1.18%	1.64[0.92,2.95]
Total (95% CI)	**11151**	**11042**		**100%**	**1.05[0.98,1.12]**
Total events: 1537 (Routine scan), 1455 (Control)					
Heterogeneity: Tau²=0; Chi²=3.53, df=4(P=0.47); I²=0%					
Test for overall effect: Z=1.3(P=0.19)					

Favours routine scan 0.1 0.2 0.5 1 2 5 10 Favours control

Analysis 1.33. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 33 Mother not satisfied with care (worried about pregnancy).

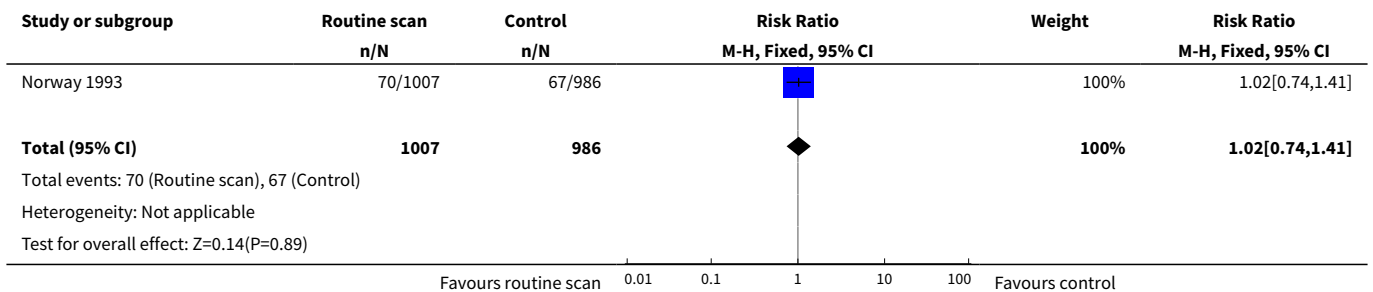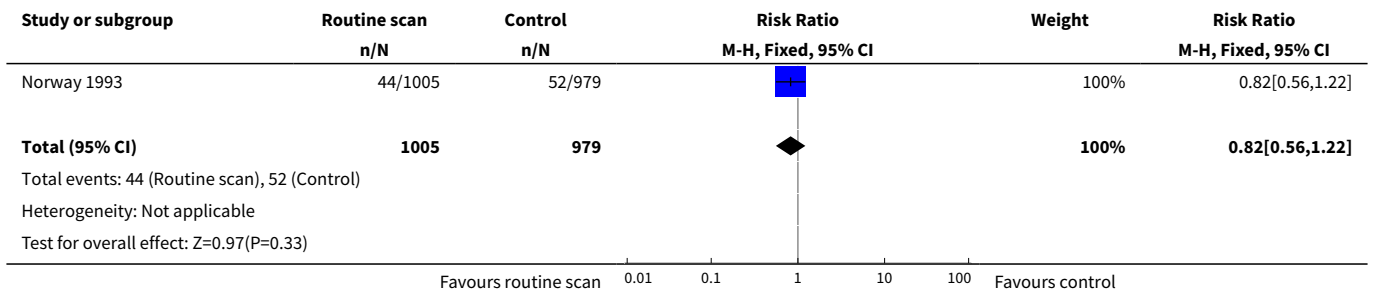| Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Adelaide 1999	98/310	128/324		100%	0.8[0.65,0.99]
Total (95% CI)	**310**	**324**		**100%**	**0.8[0.65,0.99]**
Total events: 98 (Routine scan), 128 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=2.06(P=0.04)					

Favours scan 0.01 0.1 1 10 100 Favours control

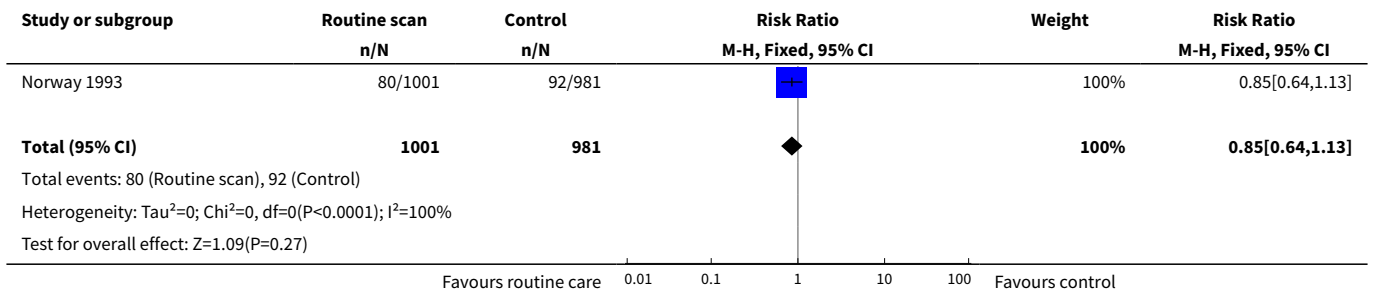### Analysis 1.34. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 34 Subgroup analysis by timing of scan: detection of multiple pregnancy by 24-26 weeks' gestation (number not detected).

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
1.34.1 Ultrasound planned before 14 weeks					
Missouri 1990	0/2	1/7		1.32%	0.89[0.05,16.36]
Subtotal (95% CI)	**2**	**7**		**1.32%**	**0.89[0.05,16.36]**
Total events: 0 (Routine scan), 1 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.08(P=0.94)					
1.34.2 Ultrasound after 14 weeks					
Alesund 1984	0/10	7/7		14.05%	0.05[0,0.73]
Helsinki 1990	0/36	9/38		14.96%	0.06[0,0.92]
London 1982	0/6	2/5		4.35%	0.17[0.01,2.92]

Favours experimental 0.01 0.1 1 10 100 Favours control

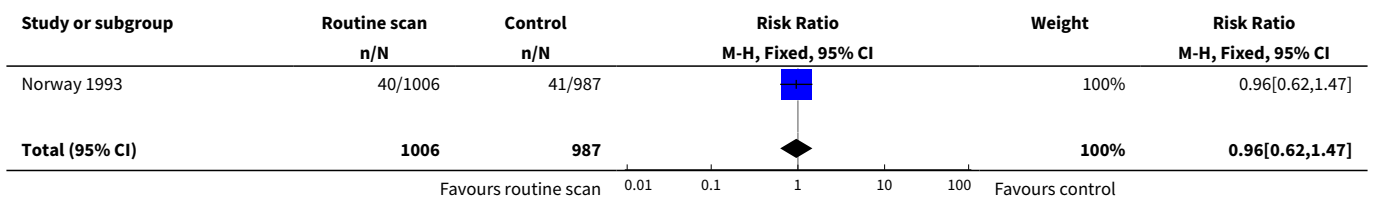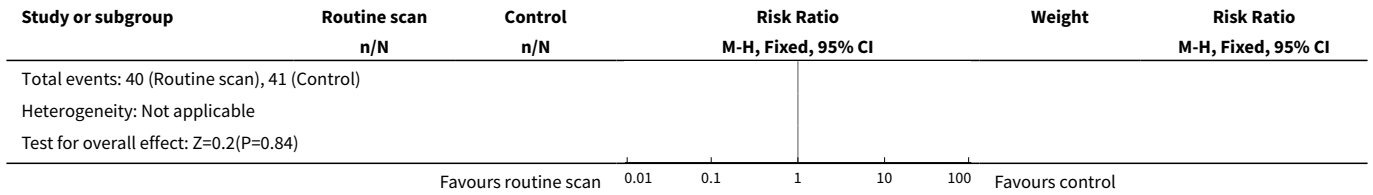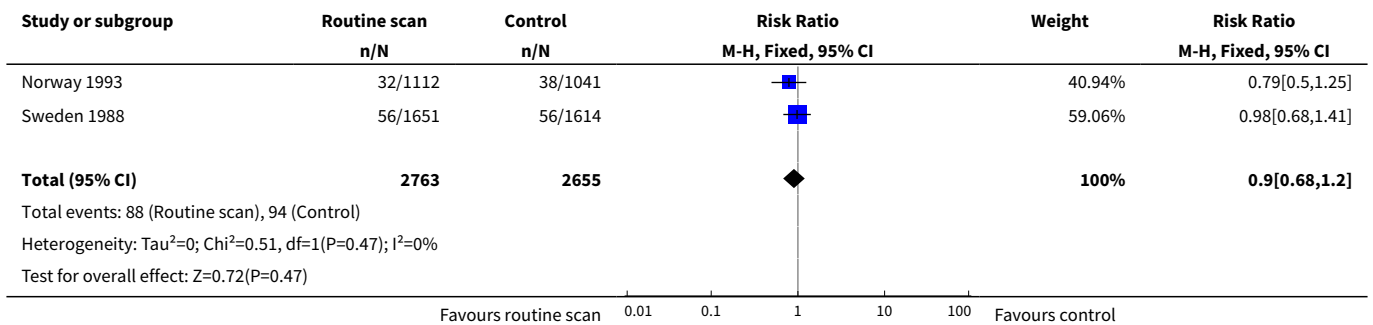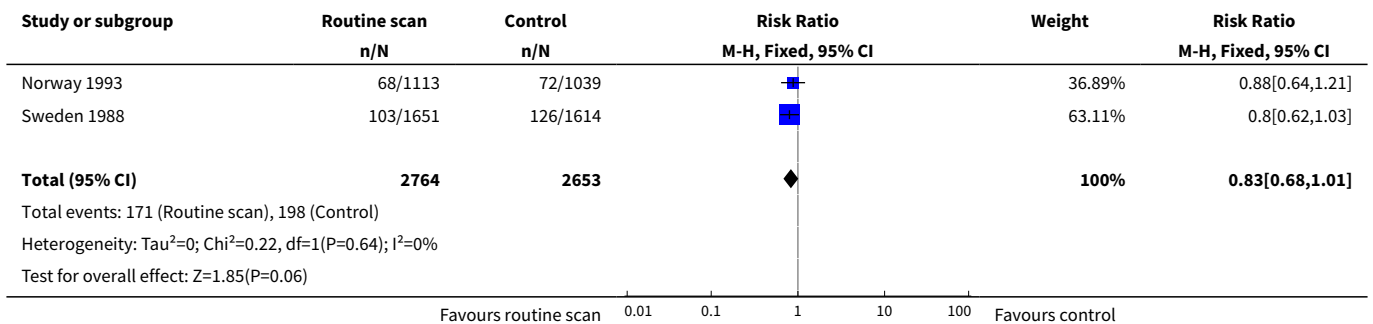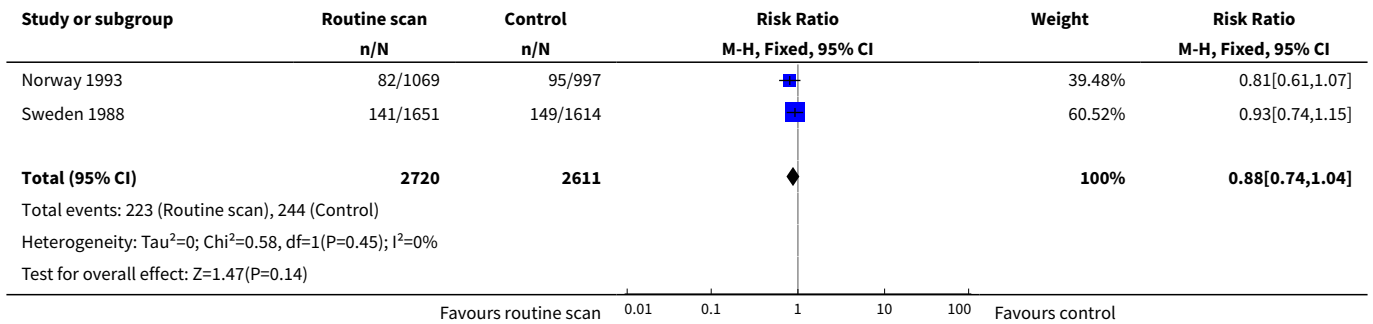| Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Fixed, 95% CI		M-H, Fixed, 95% CI
RADIUS 1993	1/68	23/61		39.22%	0.04[0.01,0.28]
Sweden 1988	0/25	13/20		24.16%	0.03[0,0.47]
Trondheim 1984	1/6	1/4		1.94%	0.67[0.06,7.85]
Subtotal (95% CI)	151	135		98.68%	0.06[0.02,0.16]
Total events: 2 (Routine scan), 55 (Control)					
Heterogeneity: Tau²=0; Chi²=4.69, df=5(P=0.46); I²=0%					
Test for overall effect: Z=5.54(P<0.0001)					
Total (95% CI)	153	142		100%	0.07[0.03,0.17]
Total events: 2 (Routine scan), 56 (Control)					
Heterogeneity: Tau²=0; Chi²=7.32, df=6(P=0.29); I²=18.09%					
Test for overall effect: Z=5.68(P<0.0001)					
Test for subgroup differences: Chi²=2.99, df=1 (P=0.08), I²=66.5%					

Favours experimental 0.01 0.1 1 10 100 Favours control

Analysis 1.35.   Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 35 Subgroup analysis: induction of labour for 'post-term' pregnancy (early and later scans).

Study or subgroup	Routine scan	Control	Risk Ratio	Weight	Risk Ratio
	n/N	n/N	M-H, Random, 95% CI		M-H, Random, 95% CI
1.35.1 Scan before 14 weeks					
Missouri 1990	28/402	31/413		16.34%	0.93[0.57,1.52]
Oxford 2006	19/233	17/230		13.77%	1.1[0.59,2.07]
Subtotal (95% CI)	635	643		30.11%	0.99[0.67,1.46]
Total events: 47 (Routine scan), 48 (Control)					
Heterogeneity: Tau²=0; Chi²=0.18, df=1(P=0.67); I²=0%					
Test for overall effect: Z=0.05(P=0.96)					
1.35.2 Scan after 14 weeks					
Alesund 1984	13/774	44/750		14.09%	0.29[0.16,0.53]
RADIUS 1993	123/7617	161/7534		21.22%	0.76[0.6,0.95]
Sweden 1988	41/2389	88/2412		18.83%	0.47[0.33,0.68]
Trondheim 1984	14/496	19/478		12.9%	0.71[0.36,1.4]
Tygerberg 1996	1/493	14/491		2.86%	0.07[0.01,0.54]
Subtotal (95% CI)	11769	11665		69.89%	0.49[0.31,0.77]
Total events: 192 (Routine scan), 326 (Control)					
Heterogeneity: Tau²=0.17; Chi²=15.78, df=4(P=0); I²=74.65%					
Test for overall effect: Z=3.08(P=0)					
Total (95% CI)	12404	12308		100%	0.61[0.42,0.88]
Total events: 239 (Routine scan), 374 (Control)					
Heterogeneity: Tau²=0.15; Chi²=21.11, df=6(P=0); I²=71.58%					
Test for overall effect: Z=2.66(P=0.01)					
Test for subgroup differences: Chi²=5.4, df=1 (P=0.02), I²=81.47%					

Favours experimental 0.01 0.1 1 10 100 Favours control

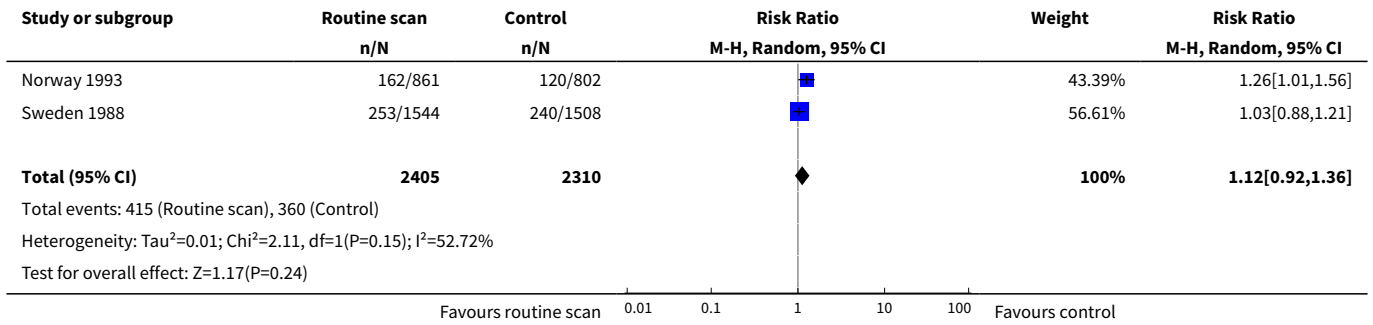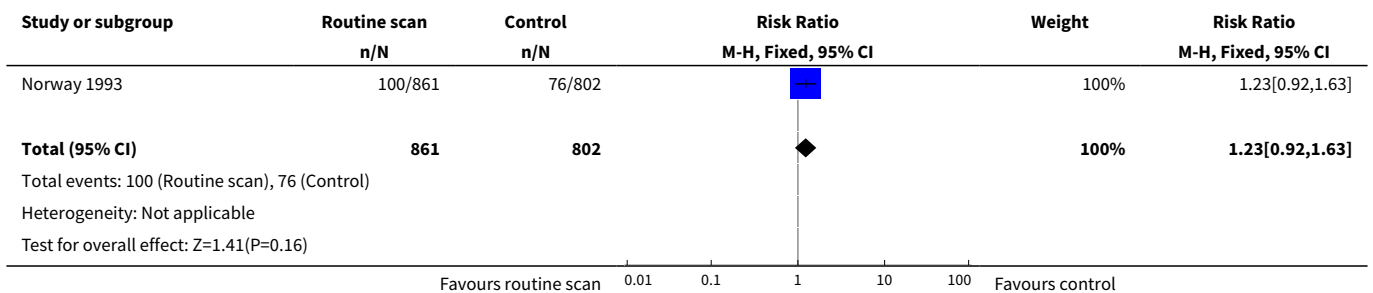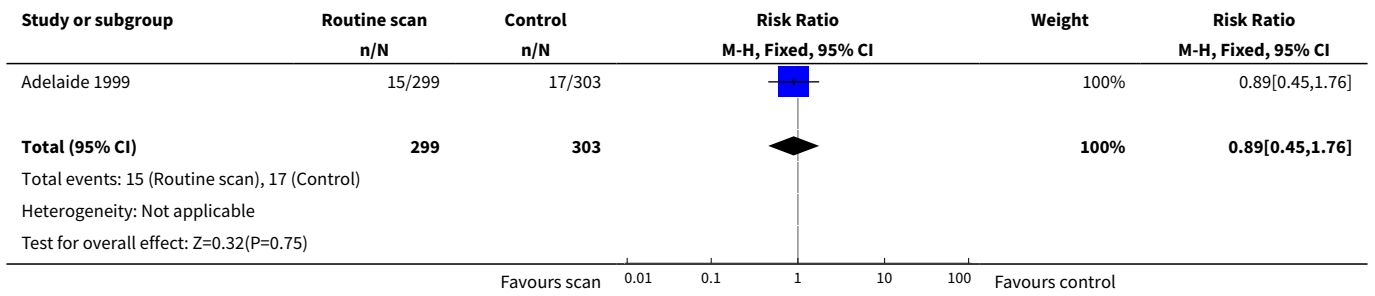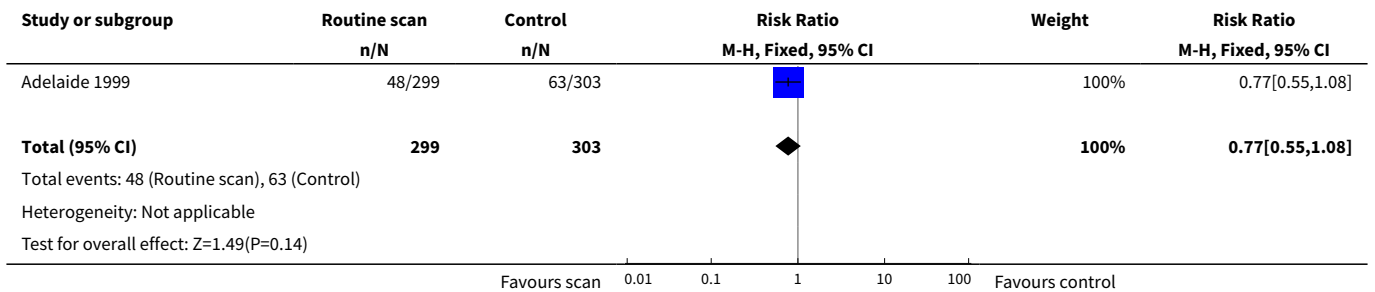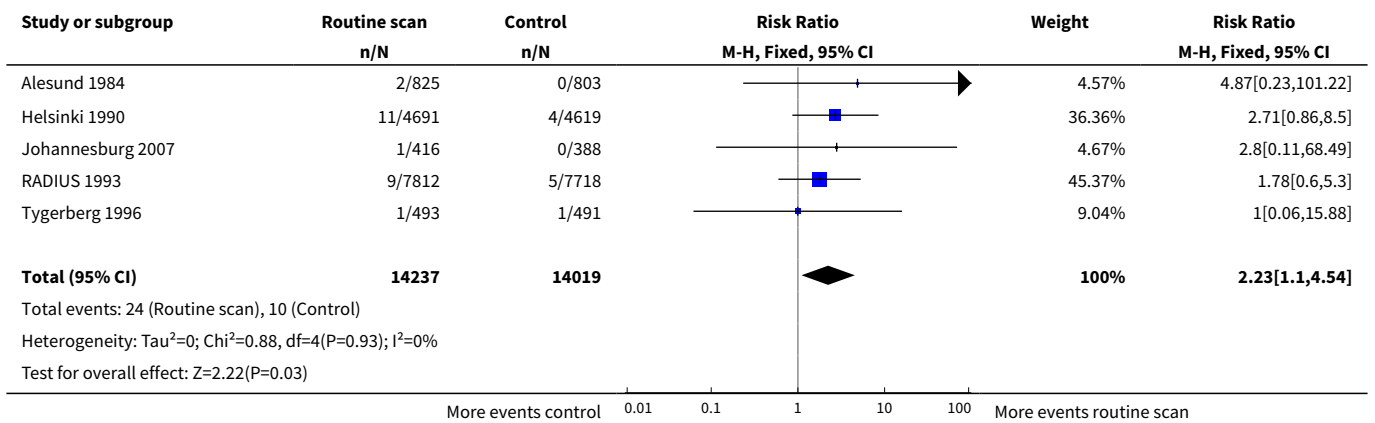### Analysis 1.36. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 36 Subgroup analysis: perinatal death (earlier and late scans).

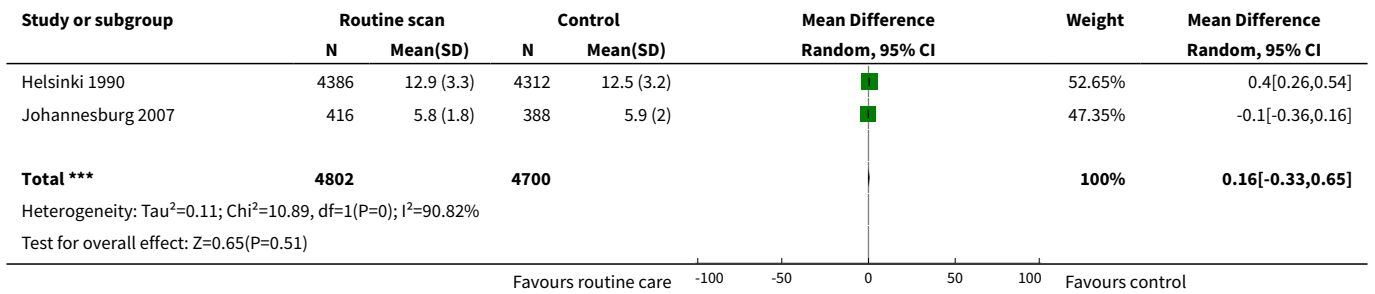| Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
1.36.1 Scan before 14 weeks					
Adelaide 1999	3/296	3/298		2.29%	1.01[0.2,4.95]
Missouri 1990	2/402	4/420		2.99%	0.52[0.1,2.84]
Subtotal (95% CI)	**698**	**718**		**5.28%**	**0.73[0.23,2.3]**
Total events: 5 (Routine scan), 7 (Control)					
Heterogeneity: Tau²=0; Chi²=0.31, df=1(P=0.58); I²=0%					
Test for overall effect: Z=0.53(P=0.59)					
1.36.2 Scan after 14 weeks					
Alesund 1984	6/794	10/765		7.79%	0.58[0.21,1.58]
Helsinki 1990	20/4422	39/4350		30.07%	0.5[0.29,0.86]
London 1982	5/537	3/536		2.3%	1.66[0.4,6.93]
RADIUS 1993	52/7685	41/7596		31.54%	1.25[0.83,1.89]
Sweden 1988	12/2415	12/2432		9.15%	1.01[0.45,2.24]
Trondheim 1984	5/503	5/482		3.91%	0.96[0.28,3.29]
Tygerberg 1996	9/496	13/494		9.96%	0.69[0.3,1.6]
Subtotal (95% CI)	**16852**	**16655**		**94.72%**	**0.87[0.68,1.13]**
Total events: 109 (Routine scan), 123 (Control)					
Heterogeneity: Tau²=0; Chi²=8.89, df=6(P=0.18); I²=32.54%					
Test for overall effect: Z=1.02(P=0.31)					
Total (95% CI)	**17550**	**17373**		**100%**	**0.87[0.68,1.11]**
Total events: 114 (Routine scan), 130 (Control)					
Heterogeneity: Tau²=0; Chi²=9.28, df=8(P=0.32); I²=13.83%					
Test for overall effect: Z=1.11(P=0.27)					
Test for subgroup differences: Chi²=0.09, df=1 (P=0.77), I²=0%					

Favours experimental 0.01 0.1 1 10 100 Favours control

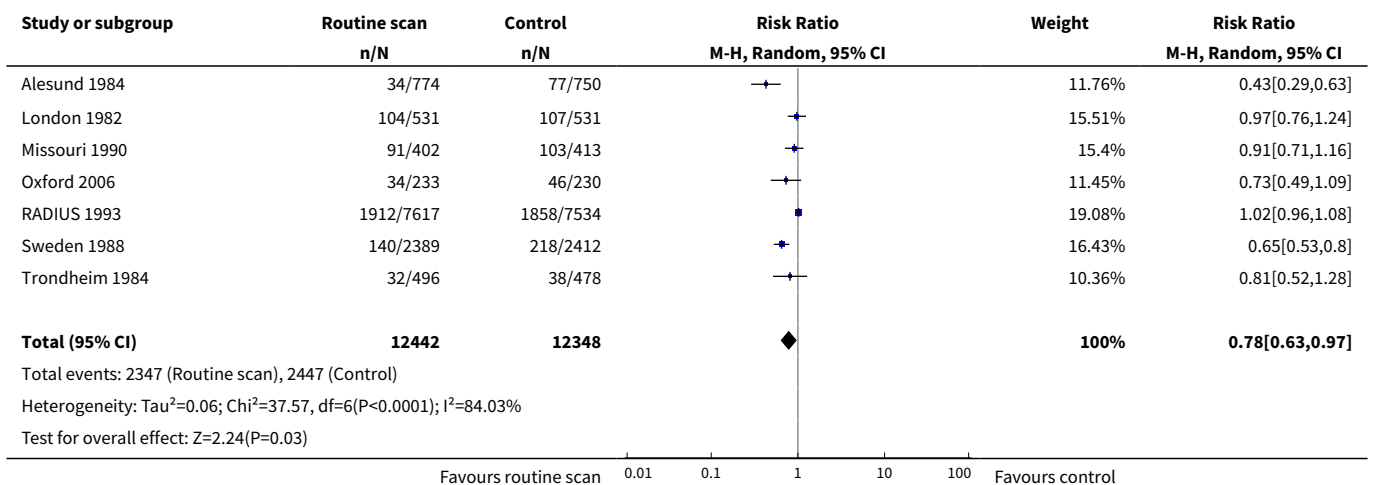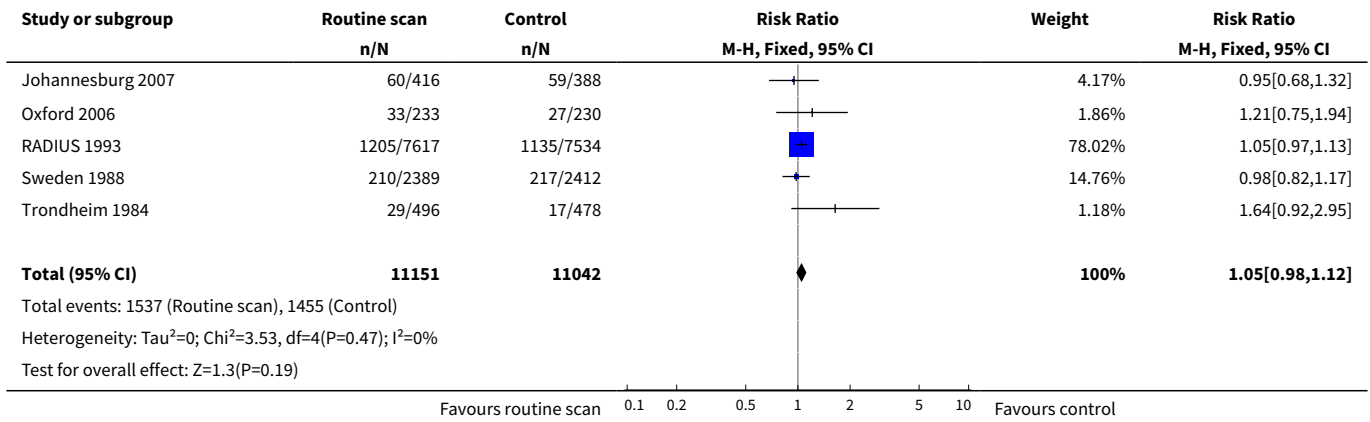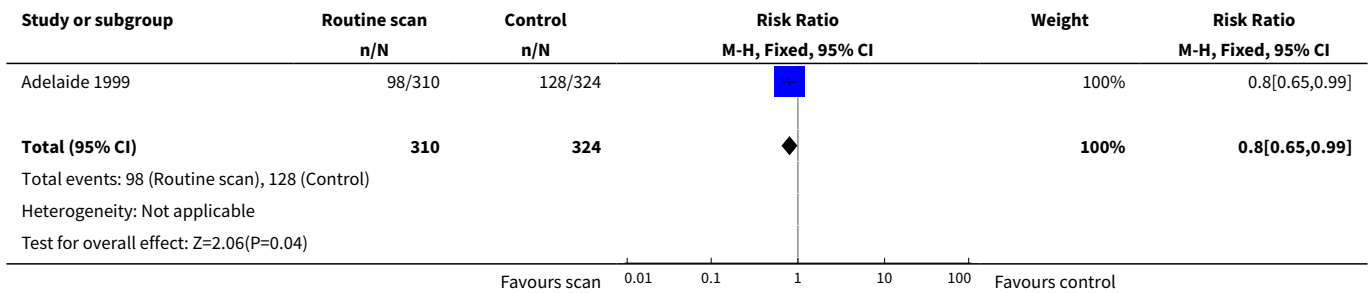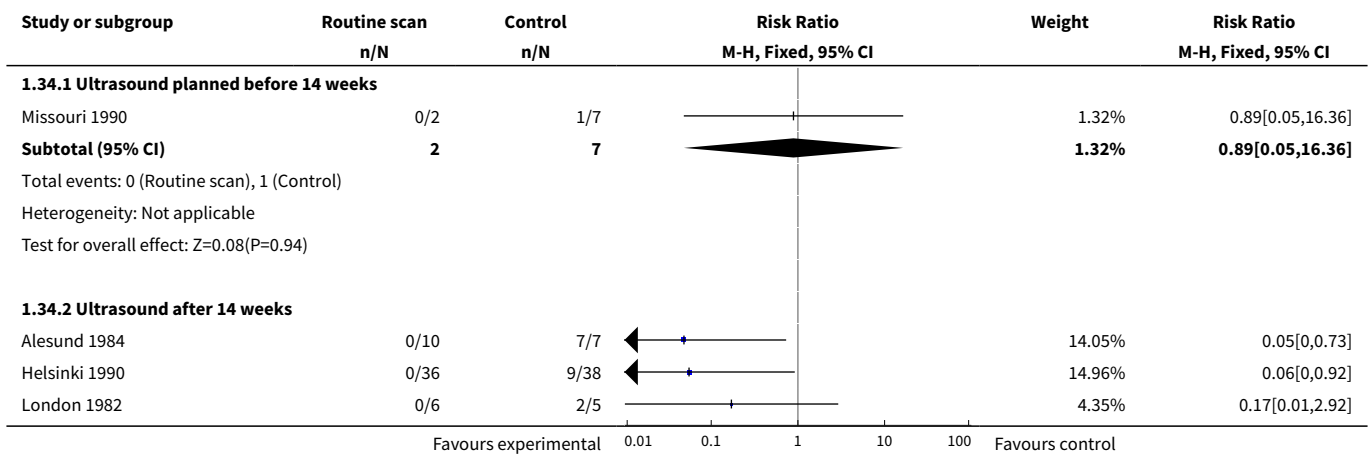### Analysis 1.37. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 37 Subgroup analysis: detection of multiple pregnancy before 24 weeks (number not detected; concealed results.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
1.37.1 Concealed results for controls					
London 1982	0/6	2/5		4.35%	0.17[0.01,2.92]
Subtotal (95% CI)	**6**	**5**		**4.35%**	**0.17[0.01,2.92]**
Total events: 0 (Routine scan), 2 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=1.22(P=0.22)					
1.37.2 Selective scan for controls					
Alesund 1984	0/10	7/7		14.05%	0.05[0,0.73]
Helsinki 1990	0/36	9/38		14.96%	0.06[0,0.92]
Missouri 1990	0/2	1/7		1.32%	0.89[0.05,16.36]
RADIUS 1993	1/68	23/61		39.22%	0.04[0.01,0.28]
Sweden 1988	0/25	13/20		24.16%	0.03[0,0.47]
Trondheim 1984	1/6	1/4		1.94%	0.67[0.06,7.85]

Favours experimental 0.01 0.1 1 10 100 Favours control

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Subtotal (95% CI)	147	137	◆	95.65%	0.07[0.02,0.17]
Total events: 2 (Routine scan), 54 (Control)					
Heterogeneity: Tau²=0; Chi²=7.13, df=5(P=0.21); I²=29.88%					
Test for overall effect: Z=5.51(P<0.0001)					
Total (95% CI)	153	142	◆	100%	0.07[0.03,0.17]
Total events: 2 (Routine scan), 56 (Control)					
Heterogeneity: Tau²=0; Chi²=7.32, df=6(P=0.29); I²=18.09%					
Test for overall effect: Z=5.68(P<0.0001)					
Test for subgroup differences: Chi²=0.4, df=1 (P=0.53), I²=0%					

Favours experimental 0.01 0.1 1 10 100 Favours control

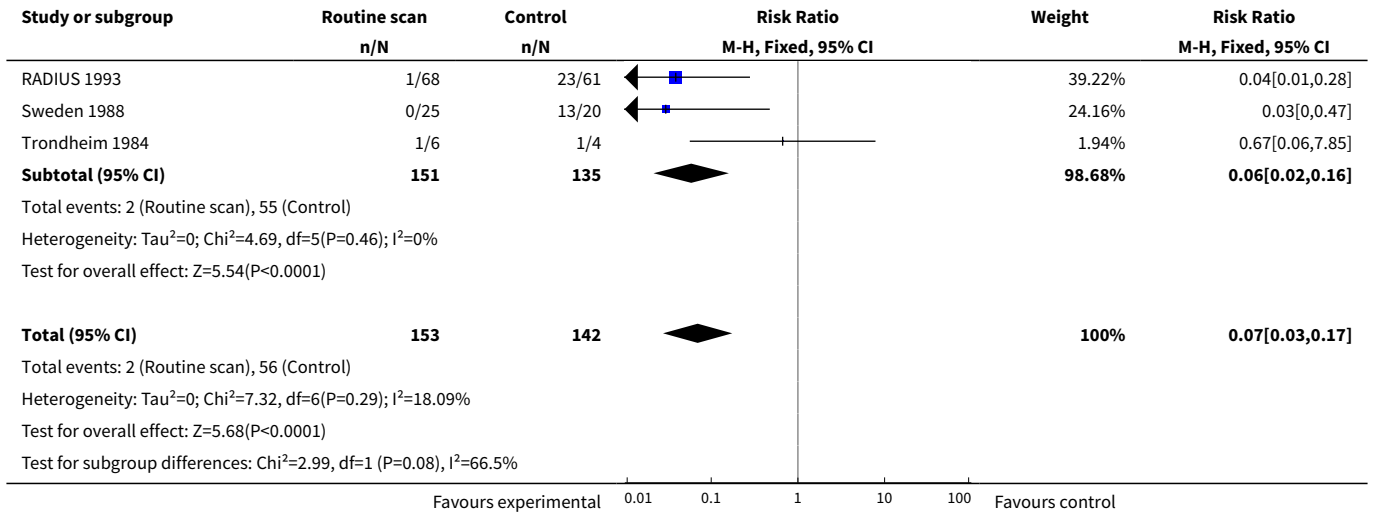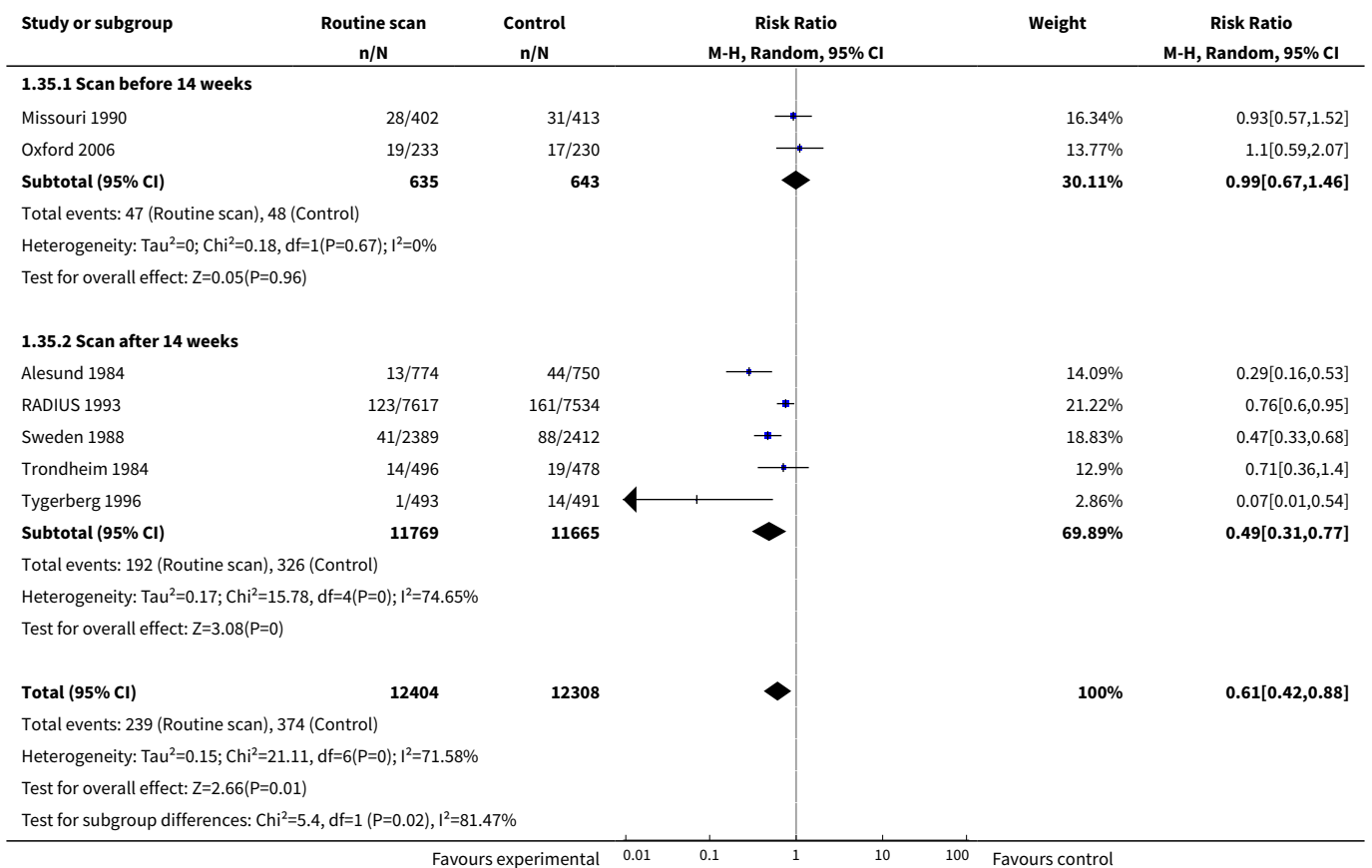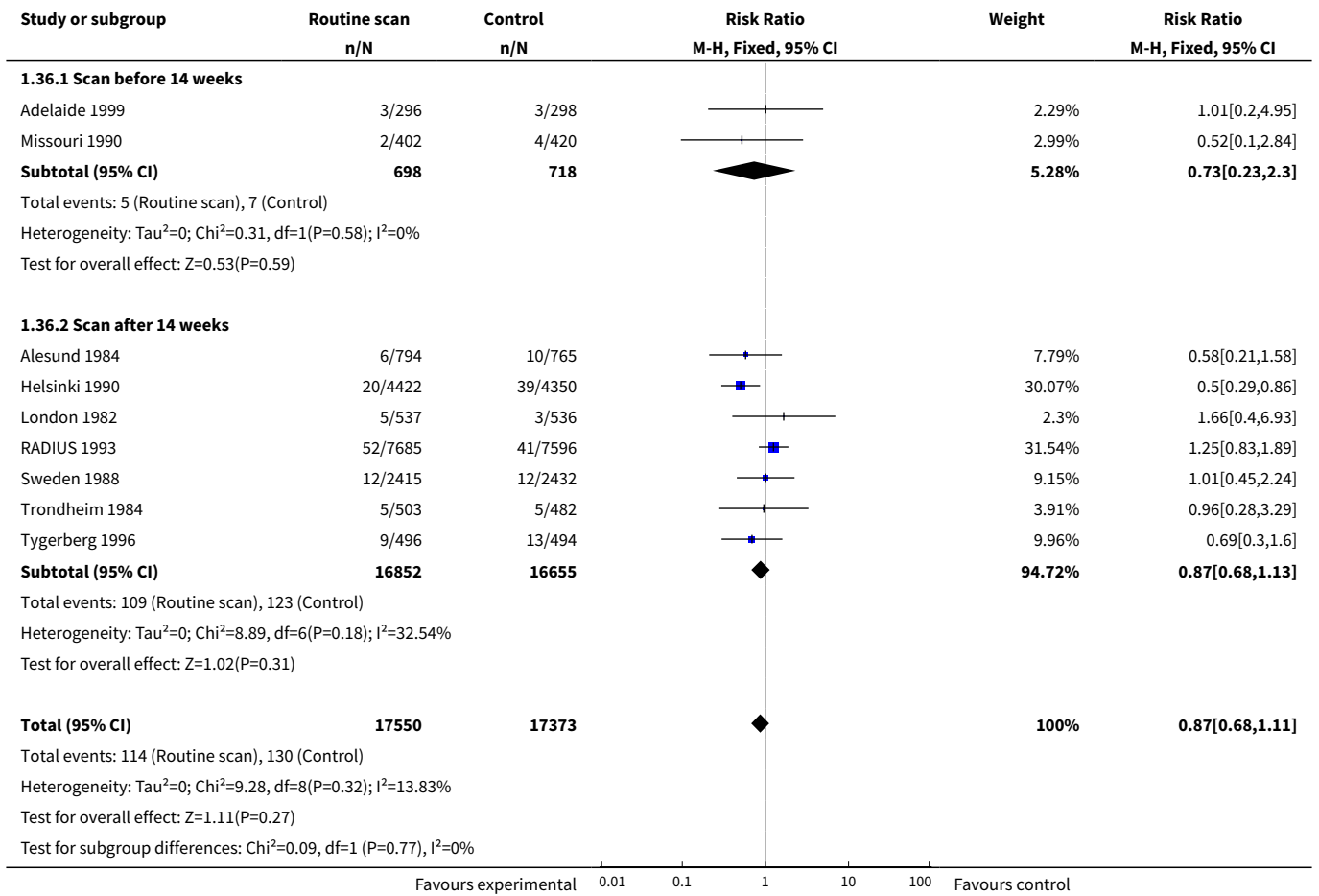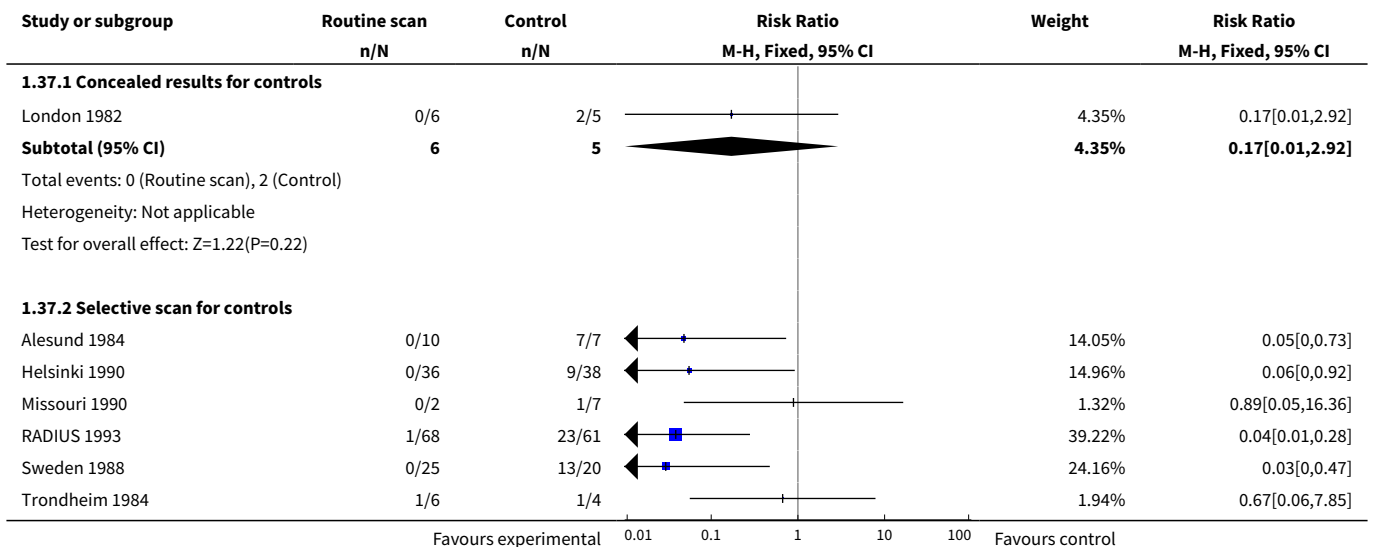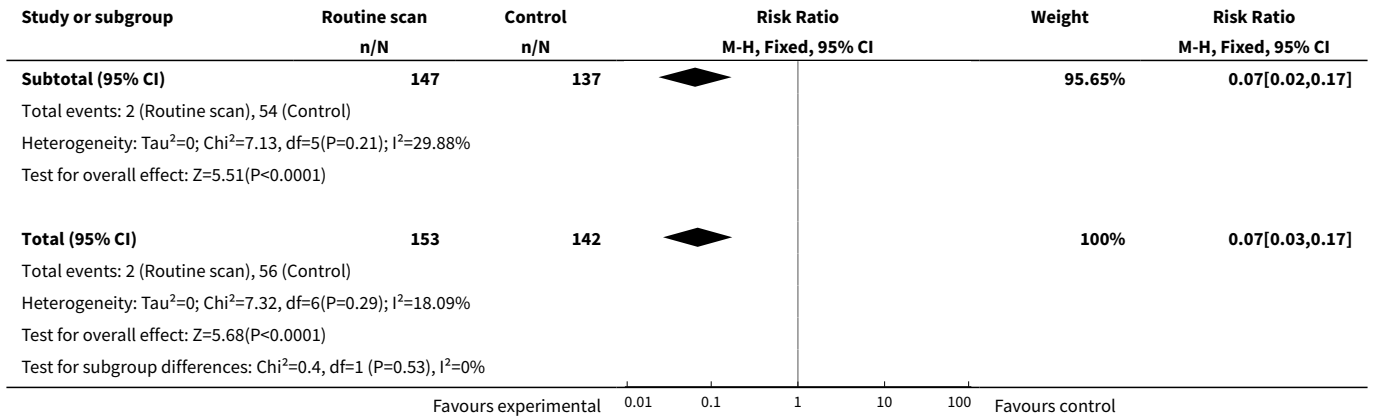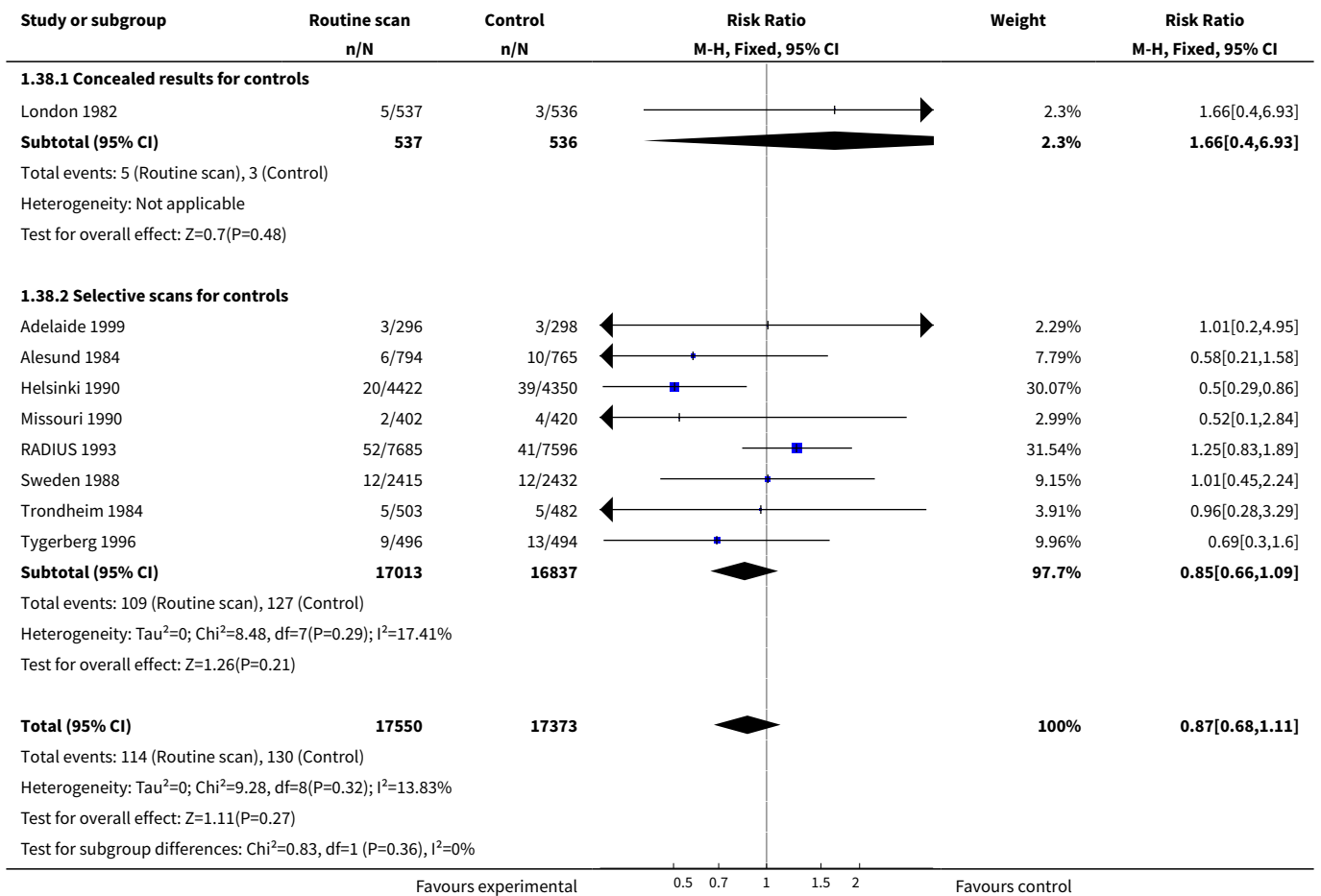### Analysis 1.38. Comparison 1 Routine/revealed versus selective/concealed ultrasound in early pregnancy, Outcome 38 Subgroup analysis: perinatal death. Concealed results.

Study or subgroup	Routine scan n/N	Control n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
1.38.1 Concealed results for controls					
London 1982	5/537	3/536	▸	2.3%	1.66[0.4,6.93]
Subtotal (95% CI)	537	536	◆	2.3%	1.66[0.4,6.93]
Total events: 5 (Routine scan), 3 (Control)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.7(P=0.48)					
1.38.2 Selective scans for controls					
Adelaide 1999	3/296	3/298	◂▸	2.29%	1.01[0.2,4.95]
Alesund 1984	6/794	10/765	◂▪	7.79%	0.58[0.21,1.58]
Helsinki 1990	20/4422	39/4350	▪	30.07%	0.5[0.29,0.86]
Missouri 1990	2/402	4/420	◂	2.99%	0.52[0.1,2.84]
RADIUS 1993	52/7685	41/7596	▪	31.54%	1.25[0.83,1.89]
Sweden 1988	12/2415	12/2432	▪	9.15%	1.01[0.45,2.24]
Trondheim 1984	5/503	5/482	◂	3.91%	0.96[0.28,3.29]
Tygerberg 1996	9/496	13/494	▪	9.96%	0.69[0.3,1.6]
Subtotal (95% CI)	17013	16837	◆	97.7%	0.85[0.66,1.09]
Total events: 109 (Routine scan), 127 (Control)					
Heterogeneity: Tau²=0; Chi²=8.48, df=7(P=0.29); I²=17.41%					
Test for overall effect: Z=1.26(P=0.21)					
Total (95% CI)	17550	17373	◆	100%	0.87[0.68,1.11]
Total events: 114 (Routine scan), 130 (Control)					
Heterogeneity: Tau²=0; Chi²=9.28, df=8(P=0.32); I²=13.83%					
Test for overall effect: Z=1.11(P=0.27)					
Test for subgroup differences: Chi²=0.83, df=1 (P=0.36), I²=0%					

Favours experimental 0.5 0.7 1 1.5 2 Favours control

WHAT'S NEW

Date	Event	Description
30 March 2015	New search has been performed	Search updated. Six new reports identified: one is a report of an ongoing trial (McClure 2014); two reports are additional reports of studies already included (Sweden 1988; Johannesburg 2007); three reports have been excluded (Georgsson 2010; Votino 2012; Schifano 2010). A 'Summary of findings' table has been incorporated.
30 March 2015	New citation required but conclusions have not changed	Conclusions not changed.

CONTRIBUTIONS OF AUTHORS

MK Whitworth (MKW) drafted the initial protocol and review. All other authors commented on subsequent drafts of the review. Data extraction was carried out by MKW, L Bricker and T Dowswell.

For the 2015 update, MKW updated the review with contributions from L Bricker and C Mullan.

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- The University of Liverpool, UK.
- Liverpool Women's NHS Foundation Trust, UK.

External sources

- National Institute for Health Research (NIHR), UK.

  TD was supported by the NIHR NHS Cochrane Collaboration Programme grant scheme award for NHS-prioritised centrally-managed, pregnancy and childbirth systematic reviews: CPGS02 (2010)
- UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), Department of Reproductive Health and Research (RHR), World Health Organization, Switzerland.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

The methods section has been updated and a 'Summary of findings' table incorporated.

INDEX TERMS

Medical Subject Headings (MeSH)

*Pregnancy, Multiple; Congenital Abnormalities [*diagnostic imaging]; Fetal Monitoring [*methods]; Gestational Age; Perinatal Mortality; Pregnancy Trimester, First; Pregnancy Trimester, Second; Ultrasonography, Prenatal [*methods]

MeSH check words

Female; Humans; Infant, Newborn; Pregnancy