A critical review of the Cochrane meta-analysis of routine late-pregnancy ultrasound

GCS Smith

Department of Obstetrics and Gynaecology, NIHR Cambridge Biomedical Research Centre, University of Cambridge, Cambridge, UK

Correspondence: Prof GCS Smith, Box 223 The Rosie Hospital, Robinson Way, Cambridge, CB2 0SW, UK. Email: gcss2@cam.ac.uk

Accepted 19 June 2020. Published Online 28 July 2020.

A Cochrane review of universal late-pregnancy ultrasound has been highly influential in guiding UK practice, concluding that it does not improve outcome. However, the meta-analysis combines trials that used diverse definitions of screen positive, were designed in the absence of high-quality data on diagnostic effectiveness and did not couple screening to an effective intervention. Moreover, even if the trials had combined a highly effective screening test with a highly effective intervention, the sample size was 15% of that required to study perinatal death. It is not known whether universal late-pregnancy ultrasound confers benefit on the mother or baby.

Keywords Meta-analysis, screening, ultrasound.

Tweetable abstract Despite >50 years of research, we do not know whether universal late-pregnancy ultrasound confers benefit on the mother or baby.

Introduction

Current UK practice

Since the first descriptions of obstetric ultrasound almost 60 years ago, it has become a core tool in modern obstetrics. In current practice in the UK, all women are offered screening in the first half of pregnancy using ultrasound, at around 12 weeks of gestation and around 20 weeks of gestation. The screening information retrieved includes accurate gestational dating; identification of multiple pregnancy; identification of other gynaecological pathology; assessment of the risk and presence of congenital anomalies, including aneuploidy; and placental localisation. Scanning after 20 weeks is not offered to all women but rather it is targeted on the basis of risk factors (e.g. previous complicated pregnancy), physical examination (e.g. suspected small for gestational age), the experience of symptoms (e.g. antepartum haemorrhage) or development of complications (e.g. pre-eclampsia). In some countries (e.g. France), all women are offered an ultrasound scan in late pregnancy. In the UK, this option has been considered by the Royal College of Obstetricians and Gynaecologists, the National Institute for Health and Care Excellence and the Department of Health, and they all decided against offering universal ultrasound, citing a Cochrane review that concluded ‘routine late-pregnancy ultrasound in low-risk or unselected populations does not confer benefit on mother or baby’.

Potential for screening using universal ultrasound in late pregnancy

Some of the features that could be detected by screening with late-pregnancy ultrasound could plausibly lead to interventions that improve outcome. Perhaps the simplest of these is the presentation of the infant. Breech presentation affects 3–4% of all pregnancies at term. Clinicians currently screen for breech presentation using physical examination, but this is known to have a sensitivity of 50–70%. Hence, a significant proportion of term labours are complicated by an undiagnosed breech presentation. In such cases, if delivery is not imminent, women would usually be offered emergency caesarean delivery. The basis for this is that vaginal breech birth is associated with a significantly increased risk of intrapartum stillbirth and neonatal death and the Term Breech Study demonstrated that this risk can be reduced by caesarean delivery. Epidemiological studies confirmed that, following the practice changes prompted by the Term Breech Study, the rates of intrapartum stillbirth and neonatal death associated with vaginal breech birth declined sharply at a population level. Hence, a combination of biological plausibility, epidemiological data and a randomised controlled trial indicate that knowing the fetal presentation near term is breech could prevent the death of the infant. It is estimated that five to ten babies a year in the UK die as the result of an undiagnosed breech presentation in labour. The mother...
Screening and intervention

Screening is a two-stage process. First, some form of assessment is made to identify women who either have occult disease, or are at risk of developing disease. Second, an intervention is applied that either treats the early-stage disease or reduces the risk of disease development in those at high risk. It follows, therefore, that a successful programme – or research trial – must address both elements of the process. Having a diagnostically effective screening test is necessary, but not sufficient, for a clinically effective programme of screening and intervention.

Diagnostic test

Ultrasound is not, itself, a diagnostic test. Rather it is an imaging modality that allows diagnostic tests to be performed. In the second half of pregnancy, ultrasound can be used to determine multiple features that could be useful in the prediction of adverse outcome including presentation; fetal biometry, including identification of fetuses that are small or large for gestational age; assessment of amniotic fluid volume; assessment of placental appearance; detection of vasa praevia; measurement of indices from Doppler flow velocimetry of maternal, placental or fetal blood vessels; assessment of fetal biological state, such as activity or fetal breathing movements; and re-screening for fetal anomalies. The timing of the scan is also relevant. The primary way that late-pregnancy ultrasound informs obstetric care in the second half of pregnancy is in relation to the timing or mode of delivery. Hence, combining studies comprising scans at different gestational ages is potentially problematic. Other complexities include the use of serial scans, the measurements and reference ranges used for fetal biometry, identifying thresholds for defining screen positives and combining ultrasound with other variables for risk assessment.

Diagnostic effectiveness is assessed using a number of statistical measures. Some of these express the effectiveness of the test at the level of the population: the sensitivity is the proportion of all women who will experience the outcome who test positive, and the specificity is the proportion of all women who will not experience the outcome who test negative. These statistics allow the likely population impact of screening to be assessed in terms of what proportion of disease might be prevented and what proportion of healthy women would be wrongly classified as high risk. However, from the perspective of the individual woman, the positive and negative predictive values (PPV and NPV, respectively) are more relevant. The PPV is the proportion of women who screen positive who subsequently experience the outcome and the NPV is the proportion of women who screen negative who remain healthy. The PPV and NPV are determined by the combined effects of two other factors: first, the woman’s prior risk of disease and, second, the proportional change in the risk associated with a positive or negative test result, i.e. the positive likelihood ratio and negative likelihood ratio, respectively.

The importance of prior risk in a screening study can be illustrated with the case of pre-eclampsia (Figure 1). If a woman has had a previous uncomplicated term labour, her risk of pre-eclampsia in future pregnancies is ~1%. Hence, even if a test has a positive likelihood ratio of 5, she still has only a ~5% chance that she will experience the disease. If a woman has had a previous preterm delivery due to pre-eclampsia, her risk of recurrence is up to 40% in the next pregnancy. Hence, even if the test has a negative likelihood ratio of 0.2, she still has a >10% risk of developing the condition. Women with no previous births are at an intermediate level of risk and it has been suggested that it is in these women that screening has the greatest potential to be clinically useful. This can be illustrated in the ASPRE trial, where 26 941 women were screened for pre-eclampsia risk and high-risk women were randomised to aspirin or placebo. Although parous women with no previous history of pre-eclampsia composed approximately half of the women screened, there were only 11 cases of preterm pre-eclampsia in such women in the screen-positive group of the ASPRE trial.

It follows, therefore, that when considering screening studies individually and when planning to combine screening studies in a meta-analysis, there are some key questions to consider in relation to the screening test. First, which of the many tests that the imaging modality of ultrasound can perform were actually conducted? Second, at what gestational age was the screening test performed? Third, what is the diagnostic effectiveness of the screening test in a low-risk or unselected population?

Intervention

As discussed above, a clinically effective programme of screening and intervention also requires that the screening test is coupled to an intervention that reduces the risk of the given outcome in high-risk women. The primary disease-modifying intervention in the third trimester of pregnancy is the decision to effect a medically indicated delivery.
Truncating the duration of pregnancy reduces the risk of events that can occur, or primarily only occur, in women who are still pregnant, such as stillbirth or pre-eclampsia. Reducing the duration of pregnancy is a benign intervention post-term, as the natural history of advancing gestational age post-term is that the risk of adverse outcomes increases. However, for the earlier period of the third trimester, the risks of many adverse outcomes for the baby decline with advancing gestational age. This is true for both short-term and long-term adverse outcomes, and the risk may continue to decline up to 40 weeks of gestation. Consequently, even early term delivery (at 37–38 weeks of gestation), although relatively benign in comparison with preterm delivery, is still associated with an increased risk of adverse outcome for the infant.

The above then leads to the crux of the problem in screening for complications in the preterm period, namely, that intervention will generally increase the risk of relatively common adverse neonatal outcomes for the baby but may reduce the risk of intrauterine fetal death leading to stillbirth. Use of umbilical artery Doppler flow velocimetry in high-risk pregnancies has been shown to result in a reduction in perinatal mortality. However, as the screening was generally performed in the context of preterm fetal growth restriction, the beneficial effect of umbilical artery Doppler imaging may also have been mediated by preventing unnecessary iatrogenic preterm delivery, as well as by promoting earlier delivery to prevent stillbirth. Furthermore, in the context of screening low-risk or unselected women early in the third trimester, the lower prior risk of complications means that abnormal scan results will be associated with a lower PPV. Hence, the ratio of false positives to true positives will be higher and it becomes more likely that the net effect of intervention is likely to be harmful.

Truncating the duration of pregnancy reduces the risk of events that can occur, or primarily only occur, in women who are still pregnant, such as stillbirth or pre-eclampsia. Reducing the duration of pregnancy is a benign intervention post-term, as the natural history of advancing gestational age post-term is that the risk of adverse outcomes increases. However, for the earlier period of the third trimester, the risks of many adverse outcomes for the baby decline with advancing gestational age. This is true for both short-term and long-term adverse outcomes, and the risk may continue to decline up to 40 weeks of gestation. Consequently, even early term delivery (at 37–38 weeks of gestation), although relatively benign in comparison with preterm delivery, is still associated with an increased risk of adverse outcome for the infant.

The above then leads to the crux of the problem in screening for complications in the preterm period, namely, that intervention will generally increase the risk of relatively common adverse neonatal outcomes for the baby but may reduce the risk of intrauterine fetal death leading to stillbirth. Use of umbilical artery Doppler flow velocimetry in high-risk pregnancies has been shown to result in a reduction in perinatal mortality. However, as the screening was generally performed in the context of preterm fetal growth restriction, the beneficial effect of umbilical artery Doppler imaging may also have been mediated by preventing unnecessary iatrogenic preterm delivery, as well as by promoting earlier delivery to prevent stillbirth. Furthermore, in the context of screening low-risk or unselected women early in the third trimester, the lower prior risk of complications means that abnormal scan results will be associated with a lower PPV. Hence, the ratio of false positives to true positives will be higher and it becomes more likely that the net effect of intervention is likely to be harmful.

For example, in the Pregnancy Outcome Prediction study, we studied an unselected population of nulliparous women in Cambridge, UK and identified 118 women at 28 weeks of gestation where a blinded ultrasound scan fulfilled the definition for early fetal growth restriction, as agreed by an international Delphi
consensus process of fetal medicine specialists. This study allowed us to determine the natural history of the diagnosis and we found that only 12 of these women actually delivered a small baby preterm and none of them experienced a stillbirth. However, the positive likelihood ratio for preterm delivery of a small baby was 17. This illustrates the problem about applying even good tests to low-risk populations.

**Design of trials**

The simplest way to conduct a trial of screening is to randomise women to being screened or not being screened. However, this approach has a number of issues. First, if the trial reports a negative result, it is difficult to determine whether this was because of failure of the screening test, failure of the intervention, or failure of both. Second, this study design requires a test with high sensitivity. If the test only identifies a minority of cases of the adverse outcome in the population, it follows that the majority of adverse outcomes in both arms of the trial were not potentially preventable by screening and intervention. Although it might seem reasonable that only tests with high sensitivity should be considered, this view does not take into account the diverse pathogenesis of many major complications of pregnancy. For example, stillbirth can be the end result of multiple different pathways. It is unlikely that a single test will be effective at identifying all fetuses at high risk of this event. Hence, for the many obstetric complications that have multiple determinants, trials of screening and intervention may be unable to demonstrate an effect if conducted using screen versus no screen. A consequence of the above is that much larger sample sizes are required for trials of screen versus no screen. Unless the power calculation for the given study is informed by a realistic assessment of the given test’s sensitivity in the population of interest, there will be a tendency for all trials, and the meta-analysis of trials, to report negative results.

The alternative study design is to screen women and assess their risk and then randomise high-risk women to having the intervention or receiving routine care (Figure 2). This study design reduces the sample size in most cases, and often very substantially (even up to 90%). The other advantage of this study design is that the individual elements of the screening programme can be assessed separately, i.e. the diagnostic effectiveness can be quantified by comparing high-risk controls with low-risk women, and the clinical effectiveness of intervention can be quantified by comparing the high-risk women randomised to intervention with high-risk controls. This issue has been reviewed in detail previously in relation to pre-eclampsia.

**Critical review of the current Cochrane meta-analysis**

The above provides a foundation for considering the Cochrane meta-analysis of universal versus selective ultrasound. Table 1 summarises some important characteristics of the eight studies included in the Cochrane review of universal ultrasound that were included in the analysis of perinatal mortality.

**Consistency of the screening test**

The table demonstrates considerable heterogeneity in the screening test employed in the eight studies. One study was confined to placental maturity. Although the rest of the studies all performed some form of fetal biometry, the nature of the fetal biometry was highly variable and multiple different equations and reference ranges were employed. Studies also variably included other parameters, such as assessment of amniotic fluid, fetal anatomy and presentation. There is an implicit assumption in the meta-analysis of perinatal mortality that is unlikely to be true, namely, that each of the definitions of screen positive had similar diagnostic effectiveness in relation to perinatal death. A further problem in relation to the screening tests was that the studies were designed and conducted without knowing the diagnostic effectiveness of the given test in the given population.

Another major question around the consistency of the screening test was the gestational age at which the scan was performed. Although some of the studies included scans near or at term, several did not. For many of the women in the meta-analysis, the finding that the fetus was in a breech presentation would be made at a gestational age when the information would be ignored. Although the review contained no specific analysis in relation to detecting breech presentation, the question of breech was briefly discussed by the authors of the systematic review. They cited a single reference in this discussion, a retrospective case series from a single UK centre published more than 5 years before the Term Breech Study. On the basis of this, the authors concluded that it ‘highlights the

---

**Figure 2.** Screening study design where randomisation follows risk assessment. The dotted lines represent the two main analyses that can be performed, quantifying the predictive ability of the screening test and quantifying the effect of the intervention in high-risk women.
uncertainty about the clinical value of routine ultrasound screening for fetal malpresentation.6

### Intervention

None of the studies listed in Table 1 coupled the screening test to a disease-modifying intervention. Some of the studies mandated a follow-up scan at a later gestational age but the majority simply revealed the result to the clinician. It is reasonable to assume that in most of the cases where the scan reported an abnormal result some form of follow up was considered. However, it is difficult to know what happened as a result of an abnormal scan result in many of the individual studies.

### Design of individual studies

All of the studies were designed as screen versus no screen, with the attendant issues for statistical power. The studies were not confined to first pregnancies. Hence, many of the pregnant women included would have been women without risk factors who had a previous uncomplicated birth. In these women, the previous risk of severe adverse events would be so low that even a very effective screening test would have a low PPV. Conversely, some studies included women with previous complicated pregnancies and high-risk women with a negative screening test could still be at significant risk of an adverse outcome.

### Statistical power

The analysis of the effect of routine ultrasound on the risk of perinatal death included 15 373 women in the routine ultrasound group and 15 302 women in the control group.6 The point estimate of the summary relative risk was 1.01 and the 95% CI was 0.67–1.54. Notwithstanding the limitations of the individual studies and pooling them, it could be argued that the lower limit of the 95% CI indicates that we can conclude with a reasonable degree of confidence that the analysis was powered to rule out the possibility that ultrasound is diagnostically effective at predicting perinatal death and that knowledge of the result of an abnormal test can lead to changes in clinical care that reduce the risk of perinatal death. However, interpretation of the 95% CI depends on an assessment of what reduction in the risk of perinatal mortality would be plausible. Taking the example of a well-established intervention, giving glucocorticoids to women in threatened preterm labour, this intervention has a summary relative risk for perinatal death of 0.72.25 Moreover, the context of steroid use is that women have attended with a specific complication and been given a specific therapy to mitigate the major risk to the infant associated with that complication. In the case of screening using ultrasound, low-risk women are being assessed by a test that will only detect risk factors for a minority of perinatal deaths and even a highly effective intervention would not be expected to prevent every adverse event. Hence, the 95% CI of the Cochrane review actually indicates a substantial degree of uncertainty, even when compared with known effective interventions applied to known high-risk groups in a context outside screening.

The lack of statistical power of the meta-analysis can also be demonstrated quantitatively. Sample size calculations for studies of screening and intervention to reduce stillbirth have been published previously (see Supplementary Figure 10 in Flodenay et al.). Assuming a background risk of perinatal death of 5 in 1000, a screening test with a 5%
screen positive rate and a positive likelihood ratio of 10, coupled to an intervention that reduces the risk of death by 40%, the rate of perinatal death would be reduced to 4 per 1000. A sample size calculation indicates that a study would need to recruit >200 000 women to have 90% power to detect this effect using standard statistical thresholds (α = 0.05, two-sided). Moreover, using these figures, the relative risk associated with screening and intervention would be 0.8. Hence, the Cochrane review had a sample size that was about 15% of the total required to study perinatal death, using this example, and the 95% CI of the result of the meta-analysis include values of relative risk that would be achieved with a highly effective screening test coupled to a highly effective intervention.

Conclusions

The Cochrane review has issues with both the primary studies and their meta-analysis. The primary studies were all designed and conducted in the absence of knowledge of the diagnostic effectiveness of the test in the given population. None of the studies coupled the screening test to disease-modifying intervention. The meta-analysis also has issues. Ultrasound is an imaging modality, not a test. The individual studies had heterogeneous definitions of screen positive and combining them into a single analysis is difficult to justify. Finally, even if the studies had all used a diagnostic test with a positive likelihood ratio of 10 and all coupled this to an intervention that reduced the risk of perinatal death by 40%, the meta-analysis only included 15% of the sample size required for adequate statistical power to study perinatal death. The correct conclusion of the meta-analysis is that it is not known whether late-pregnancy ultrasound in low-risk or unselected populations confers benefit on mother or baby. In the case of screening for breech presentation near term, the totality of the evidence quite strongly indicates that it is likely to be both clinically effective and cost-effective.7

The Cochrane review concluded that ‘routine late-pregnancy ultrasound in low-risk or unselected populations does not confer benefit on mother or baby’. This is a positive statement. It indicates that the answer to the question ‘does late-pregnancy ultrasound in low-risk or unselected populations confer benefit on mother or baby?’ is known and that the answer is ‘no’. The conflation of the absence of evidence with the presence of negative evidence is the basis for what has been described as Evidence Based Medicine’s six dangerous words: ‘there is no evidence to suggest that...’ 27 Given the above, my view is that we do not know whether universal ultrasound in late pregnancy confers benefit on the mother or baby. It might do so, and it might not. Moreover, it may have different effects in relation to the different diagnostic tests that ultrasound can perform and the different stages in the third trimester of pregnancy when these tests might be performed. It is quite remarkable that more than 50 years after the first description of obstetric ultrasound we still do not know whether using it routinely to screen for the common complications in late pregnancy results in net benefit.

Disclosure of interests

GS receives/has received research support from Roche, GSK, Illumina & Sera Prognostics. GS has been paid to attend an advisory board by GSK and has been paid to speak at a meeting by Roche, has acted as a paid consultant to GSK and is paid as a member of a GSK Data Safety Monitoring Committee. GS is named as one of three inventors in a novel biomarker test for fetal growth filed by Cambridge Enterprise. A completed disclosure of interests form is available to view online as supporting information.

Contribution to authorship

All elements of this paper were the sole work of GS.

Details of ethics approval

Not required.

Funding

None.

References

1 Donald I. “Sonar”: a new diagnostic echo-sounding technique in obstetrics and gynaecology. Proc R Soc Med 1962;55:637–8.

2 Monier I, Blondel B, Ego A, Kaminiski M, Goffinet F, Zeitlin J. Poor effectiveness of antenatal detection of fetal growth restriction and consequences for obstetric management and neonatal outcomes: a French national study. BJOG 2015;122:518–27.

3 Royal College of Obstetrics & Gynaecologists. Green-top Guideline No. 31: The investigation and management of the small for gestational age fetus. London: Royal College of Obstetricians & Gynaecologists, 2013.

4 National Collaborating Centre for Women’s and Children’s Health. Antenatal Care – Routine Care for the Healthy Pregnant Women. London: Royal College of Obstetricians and Gynaecologists; 2008.

5 National Health Service Sins. Saving Babies’ Lives Care Bundle Version 2. London: NHS England; 2019.

6 Bricker L, Medley N, Pratt JI. Routine ultrasound in late pregnancy (after 24 weeks’ gestation). Cochrane Database Syst Rev 2015;6:CD001451.

7 Wastiund D, Moraitis AA, Dacey A, Sovio U, Wilson ECF, Smith GCS. Screening for breech presentation using universal late-pregnancy ultrasonography: a prospective cohort study and cost effectiveness analysis. PLoS Med 2019;16:e1002778.

8 Thorpe-Beeston JG, Banfield PJ, Saunders NJ. Outcome of breech delivery at term. BMJ 1992;305:746–7.

9 Hannah ME, Hannah WJ, Hewson SA, Hodnett ED, Saigal S, Willan AR. Planned caesarean section versus planned vaginal birth for breech presentation at term: a randomised multicentre trial. Term Breech Trial Collaborative Group. Lancet 2000;356:1375–83.
Rietberg CC, Elferink-Stinkens PM, Visser GH. The effect of the Term Breech Trial on medical intervention behaviour and neonatal outcome in The Netherlands: an analysis of 35,453 term breech infants. BJOG 2005;112:205–9.

11 Pasupathy D, Wood AM, Pell JP, Fleming M, Smith GC. Time trend in the risk of delivery-related perinatal and neonatal death associated with breech presentation at term. Int J Epidemiol 2009;38:490–8.

12 Hernandez-Diaz S, Toh S, Cnattingius S. Risk of pre-eclampsia in first and subsequent pregnancies: prospective cohort study. BMJ 2009;338:b2255.

13 van Oostwaard MF, Langenveld J, Schuit E, et al. Recurrence of hypertensive disorders of pregnancy: an individual patient data metaanalysis. Am J Obstet Gynecol 2015;212:624–17.

14 Smith GC. Researching new methods of screening for adverse pregnancy outcome: lessons from pre-eclampsia. PLoS Med 2012;9:e1001274.

15 Rolnik DL, Wright D, Poon LC, et al. Aspirin versus placebo in pregnancies at high risk for preterm preeclampsia. N Engl J Med 2017;377:613–22.

16 Smith GC. Life-table analysis of the risk of perinatal death at term and post term in singleton pregnancies. Am J Obstet Gynecol 2001;184:489–96.

17 Shea KM, Wilcox AJ, Little RE. Postterm delivery: a challenge for epidemiologic research. Epidemiology 1998;9:199–204.

18 MacKay DF, Wood R, King A, et al. Educational outcomes following breech delivery: a record-linkage study of 456 947 children. Int J Epidemiol 2015;44:209–17.

19 Spong CY, Mercer BM, D’alton M, Kilpatrick S, Blackwell S, Saade G. Timing of indicated late-preterm and early-term birth. Obstet Gynecol 2011;118(2 Pt 1):323

20 Afifi AC, Stampalia T, Dowssett T. Fetal and umbilical Doppler ultrasound in high-risk pregnancies. Cochrane Database Syst Rev 2017;6:CD007529.

21 Gaccioli FSU, Cook E, Hund M, Charnock-Jones DS, Smith GCS. Screening for fetal growth restriction using ultrasound and the sFLTI/PlGF ratio in nulliparous women: a prospective cohort study. Lancet Child Adolesc Health 2018;2:569–81.

22 Gordijn SJ, Beune IM, Thilaganathan B, et al. Consensus definition for placental fetal growth restriction: a Delphi procedure. Ultrasound Obstet Gynecol 2016;48:333–9.

23 Smith GC, Frents RC. Stillbirth. Lancet 2007;370:1715–25.

24 Nwosu EC, Walkinshaw S, Chia P, Manasse PR, Atlay RD. Undiagnosed breech. Br J Obstet Gynaecol 1993;100:531–5.

25 Roberts D, Brown J, Medley N, Dalziel SR. Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. Cochrane Database Syst Rev 2017;3:CD004454.

26 Flenady V, Wojcieszek AM, Middleton P, et al. Stillbirths: recall to action in high-income countries. Lancet 2016;387:691–702.

27 Braithwaite RS. A piece of my mind. EBM’s six dangerous words. JAMA 2013;310:2149–50.

28 Eik-Nes SH, Salvesen KA, Okland O, Vatten LI. Routine ultrasound fetal examination in pregnancy: the ‘Alesund’ randomized controlled trial. Ultrasound Obstet Gynecol 2000;15:473–8.

29 Mckenna D, Tharmaratnam S, Mahsud S, Bailie C, Harper A, Dornan J. A randomized trial using ultrasound to identify the high-risk fetus in a low-risk population. Obstet Gynecol 2003;101:626–32.

30 Neilson JP, Munjanja SP, Whitley CR. Screening for small for dates fetuses: a controlled trial. Br Med J (Clin Res Ed) 1984;289:1179–82.

31 Duff GB. A randomized controlled trial in a hospital population of ultrasound measurement screening for the small for dates baby. Aust N Z J Obstet Gynaecol 1993;33:374–8.

32 Proud J, Grant AM. Third trimester placental grading by ultrasonography as a test of fetal wellbeing. Br Med J (Clin Res Ed) 1987;294:1641–4.

33 Ewigman BG, Crane JP, Frigoletto FD, LeFevre ML, Bain RP, McNellis D. Effect of prenatal ultrasound screening on perinatal outcome. RADIUS Study Group. N Engl J Med 1993;329:821–7.

34 Skrastad RD, Eik-Nes SH, Sviggum O, et al. A randomized controlled trial of third-trimester routine ultrasound in a non-selected population. Acta Obstet Gynecol Scand 2013;92:1353–60.

35 Bakkeiteg LS, Eik-Nes SH, Jacobsen G, et al. Randomised controlled trial of ultrasonographic screening in pregnancy. Lancet 1984;2(8396):207–11.