## Seeing the unseen

*The importance of prenatal screening*

Lugthart, M.A.

**Publication date**
2024

# CHAPTER 7

M. BUIJTENDIJK
H. SHAH
M.A. LUGTHART
Y. DAWOOD
J. LIMPENS
B.S. DE BAKKER
M.J.B. VAN DEN HOFF
M.G. LEEFLANG
E. PAJKRT

# DIAGNOSTIC ACCURACY OF ULTRASOUND SCREENING FOR FETAL STRUCTURAL ABNORMALITIES DURING THE FIRST AND SECOND TRIMESTER OF PREGNANCY IN *low-risk and unselected populations*

COCHRANE DATABASE OF SYSTEMATIC REVIEWS 2021, ISSUE 7.

## Abstract

### Objectives

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

The main objectives of this review are to assess the diagnostic accuracy of first- and second-trimester fetal anomaly screening in low-risk pregnant women, and to compare overall performance of single- and two-stage screening approaches with regards to the number of cases detected before birth, as well as the proportion of false positive diagnoses.

### Secondary objectives

Our secondary objective is to identify potential sources of heterogeneity between studies by investigating the effect of clinical factors (i.e. prior testing, factors related to the conduct of the ultrasound examination, and healthcare setting of the study) and methodological factors (i.e. related to the study design, such as methods used for postnatal ascertainment of cases) on screening performance.

## Background

Each year, nearly four million infants worldwide — three per cent of all live births — are affected by congenital anomalies.[1,2] Congenital anomalies comprise a broad spectrum of abnormalities of prenatal origin which can cause severe long-term morbidity, disability, or, in the very worst cases, be fatal.[3] In high-income countries, an estimated 20% to 29% of deaths under the age of five result from congenital malformations.[4-6] More than half of these deaths occur during the neonatal period.[6] Prenatal ultrasound allows for the visualisation of the fetus and fetal anatomy early in pregnancy, and is widely used to detect structural anomalies before birth. Identification of anomalies through ultrasound imaging requires systematic visualisation of fetal anatomical structures, using standardised imaging planes. Traditionally, examination of the fetal anatomy is performed during the second trimester of pregnancy, between 18 and 22 weeks' gestation. The aim of this scan is to provide accurate diagnostic information for the delivery of optimised antenatal care; improve safety of birth; and enable the delivery of appropriate and potentially life-saving care without delay.[7,8] Moreover, it provides parents an opportunity to obtain multidisciplinary counselling about their child's prognosis and to discuss their management options. These may include prenatal intervention, postnatal treatment, neonatal palliative care, or elective pregnancy termination, the latter being dependent on the legal status of termination of pregnancy for fetal abnormalities in their country of residence.[7,9,10]

Before ultrasound examination is undertaken, expectant parents need to have accurate information about the nature and purpose of the test and should have a clear understanding about the reliability of the test results.[11,12] The accuracy of second-trimester screening for the detection of fetal anomalies was comprehensively reviewed in a Health Technology Assessment (HTA) in 2000 by Bricker and colleagues[13], and was reassessed in 2008 in a National institute for Health and Care Excellence (NICE) guideline.[7] Based on the HTA, the overall sensitivity of second-trimester screening was 44.7%, but varied from 15.0% to 85.3% between studies. In the 2008 update by NICE, the overall sensitivity was 35.4%, ranging from 15.0% to 92.9%. The exact causes of these extensive differences between study results are not easy to discern. Both reviews mentioned factors that may contribute to the observed heterogeneity, such as the type of healthcare setting a study is performed in, completeness of postnatal ascertainment of cases, and the types of anomalies included in studies. However, the impact of potential sources of heterogeneity on observed test accuracy estimates were not formally explored.

Since the publication of the HTA and NICE guideline, additional studies have become available, and several factors that may have a substantial impact on the detection of fetal anomalies have changed, necessitating an updated review of the literature. For instance, ultrasound technology and image resolution have improved tremendously. Furthermore, screening protocols have evolved over the years. For example, the more widespread inclusion of mandatory evaluation of the cardiac outflow tracts in ultrasound screening protocols may have a significant impact on prenatal detection of congenital

heart defects.[14] Simultaneously, many countries have started offering routine first-trimester screening for fetal anomalies as part of their public health programmes. Examples among European countries include France, Italy, Spain, and Switzerland.[9] Implementation of routine first-trimester screening for fetal anomalies changes the clinical pathway and may have an impact on overall screening performance with regards to the total number of cases detected before birth, and the number of false positive diagnoses. The accuracy of first-trimester screening for fetal anomalies has been evaluated previously in different populations.[15-17] However, first-trimester anomaly screening is a rapidly evolving field and several new studies have been published on this topic. Moreover, the impact of first-trimester screening on overall screening performance, and its role in the clinical pathway, need to be determined.

In this Cochrane Review, we will determine the diagnostic accuracy of first- and second-trimester screening for fetal structural anomalies. Accuracy of prenatal screening for fetal chromosomal abnormalities will not be determined as it is beyond the scope of this review and has been explored in detail in several Cochrane Reviews.[18-21] We will consider first- and second-trimester anomaly screening both as individual tests and when combined, to evaluate their respective roles in the clinical pathway and to assess the impact of first-trimester fetal anomaly screening on overall screening performance. As the majority of severe structural anomalies occur in pregnancies without known or pre-existing risk factors[3,22-24], we will focus on low-risk and unselected populations.

### Target condition being diagnosed

The target conditions of interest are fetal structural anomalies identifiable on ultrasound imaging. Structural anomalies are defined as anatomical malformations resulting from abnormal development of organs or body parts, which can result in the structure not being formed, being partially formed, or being formed in an abnormal fashion. These anomalies can occur in isolation (i.e. single defects), representing 75% of cases, or as a group of anomalies (i.e. multiple defects).[25] A group of defects may form part of a well-described association or syndrome, however the majority of cases occur without a known relation to one another, and are referred to as "multiple congenital anomalies". [25] Anatomical defects can occur in relation with chromosomal or genetic disease, or result from exposure to certain environmental factors, including exposure to teratogenic agents (e.g. isotretinoin, barbiturates, or radiation), nutrient deficiencies (e.g. folic acid deficiency), metabolic factors (e.g. uncontrolled diabetes), infections (e.g. toxoplasmosis, rubella), or due to mechanical problems (e.g. amniotic band constrictions, severe oligohydramnios).[3,23] However, in over 50% of cases the cause of the defect remains unclear.[3]

### Index test(s)

We will examine the diagnostic accuracy of routine first- and second-trimester ultrasound screening for the detection of fetal structural anomalies in low-risk pregnant women. First-trimester screening

will be defined as a routine ultrasound scan performed from 11 weeks and 0 days' gestation (11 + 0) to 13 weeks and six days' gestation (13 + 6), and includes examination of fetal anatomy, in accordance with international guidelines for the performance of first-trimester fetal anomaly screening.[9,26-28] Second-trimester screening will be defined as a routine ultrasound scan performed between 18 weeks and 0 days' (18 + 0) and 23 weeks and six days' (23 + 6) gestation. For this review, we have chosen to set the upper limit of the second-trimester scan at 23 + 6 weeks' gestation as this gestation is widely regarded as the limit of viability and the legal limit for pregnancy termination in countries where termination of pregnancy is restricted.[9,10]

A routine ultrasound scan during which the fetal anatomy is evaluated to screen for anomalies, is typically referred to as a fetal anomaly scan. It is important to distinguish between a routine fetal anomaly scan and a targeted or detailed ultrasound examination. A routine fetal anomaly scan is usually performed by an obstetrician or (obstetric) sonographer and is generally offered to all pregnant women as part of routine prenatal care. Women who receive this test have no known risk factors for fetal structural anomalies. A targeted or detailed ultrasound examination is used to evaluate women who are identified in advance to be at increased risk for fetal anomalies.[29] It may also be used as a diagnostic follow-up test in case of an anomaly suspected or detected at a routine ultrasound examination.[30] Detailed or targeted ultrasound examinations typically use high-quality equipment and are performed by a fetal medicine specialist or a specially trained sonographer, and they include a more extensive examination of fetal anatomical structures.[29] In this review we will only evaluate the accuracy of routine ultrasound examination in a screening setting.



## Clinical pathway

### Prior test(s)

During their first visit to an antenatal clinic, women are classified as being at high or low risk for fetal anomalies, based on assessment of their medical and obstetric history and other risk factors for fetal anomaly such as exposure to teratogens or a fetal anomaly being diagnosed in a previous pregnancy. Further risk stratification may occur if women choose to enrol in a fetal aneuploidy screening programme, which aims to identify women at increased risk of carrying a fetus with Down's syndrome (trisomy 21), Edwards' syndrome (trisomy 18), or Patau syndrome (trisomy 13). The main screening tests for these conditions include the first-trimester combined test (which relies on measurement of nuchal translucency (NT) by ultrasound scan and serum protein markers at 11 to 13 + 6 weeks' gestation), and non-invasive prenatal testing (involving analysis of fetal DNA fragments in maternal blood from 10 weeks' gestation).[31] Fetuses with an increased NT may also be detected at a routine first-trimester anomaly scan. An increased NT is considered a marker for chromosomal as well as structural abnormalities, in particular cardiac defects.[32,33] Increased NT is therefore considered an important indication for further detailed anatomical examination, which may include referral for targeted fetal echocardiography.[29,34]

## Role of index test(s)

Currently, there are two main approaches to ultrasound screening for fetal anomalies in women without known risk factors: a single routine fetal anomaly scan, performed during the second trimester of pregnancy; or two anomaly scans to screen for fetal anomalies, performed during the first and second trimester of pregnancy. In the following sections, we will refer to these screening strategies as single- and two-stage approaches, respectively. The potential pathways and roles of the index tests (first- and second-trimester screening for fetal anomalies in single- and two-stage approaches) are illustrated in Figure 1.

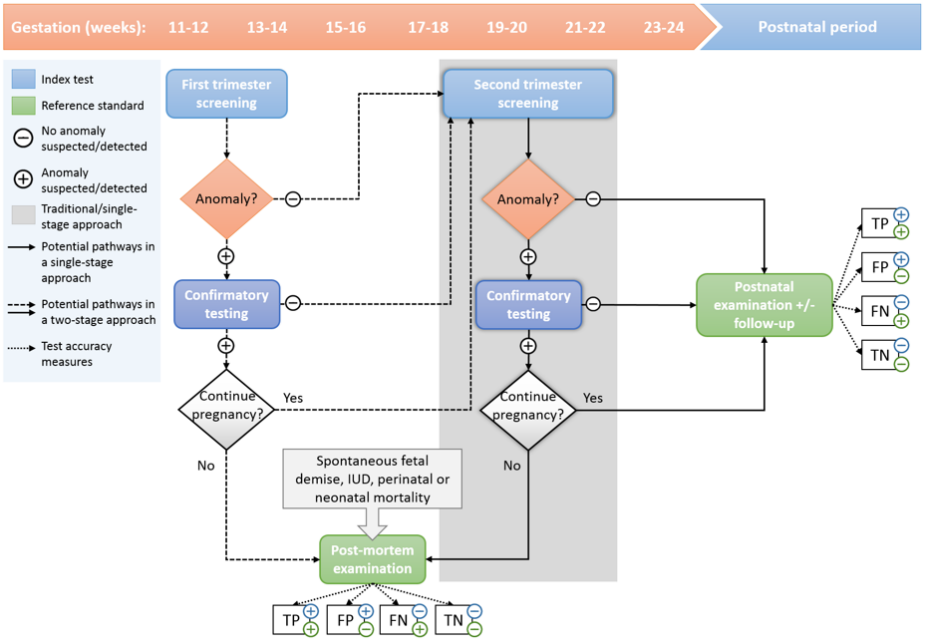


**Figure 1.** Clinical pathways in single- and two-stage screening approaches. The horizontal arrow indicates the timeline with pregnancy duration in gestational weeks and the postnatal period, respectively. In a single-stage approach (grey shaded area, solid black arrows), women are offered a fetal anomaly scan during the second trimester of pregnancy (at 18 to 22 weeks' gestation). In a two-stage approach, women are offered a fetal anomaly scan during the second trimester of pregnancy (at 18 to 22 weeks' gestation) and an additional early anomaly scan during the first trimester of pregnancy (at 11 to 13 weeks' gestation). Abbreviations: IUD: intra-uterine death; TP: true positive test result (anomaly suspected/detected on ultrasound imaging and confirmed after birth); FP: false positive test result (anomaly suspected/detected on ultrasound imaging, but not confirmed after birth); TN: true negative test result (no anomaly suspected/detected on ultrasound imaging and no anomaly found after birth); FN: false negative test result (an anomaly was found after birth, which was not detected on ultrasound imaging).

## Second-trimester screening for fetal structural anomalies: single-stage approach

In a single-stage approach, a routine second-trimester anomaly scan is the primary screening test used to detect fetal structural anomalies (see Figure 1 – grey shaded area, solid black arrows). Generally, this scan is performed between 18 and 22 weeks' gestation. If scan results are abnormal (i.e. one or more anomalies are suspected or detected), women are referred to a specialist centre

for further investigations. This may include an advanced ultrasound scan to re-evaluate or confirm the suspected anomaly, parental blood testing, invasive prenatal testing, or a combination of these.[30] Depending on the final diagnosis, parents may opt to continue or terminate the pregnancy. When parents choose to continue the pregnancy, prenatal, intrapartum, and postnatal care will be provided according to the usual obstetric standards. Postnatal care may involve additional diagnostic testing to confirm or further specify the antenatally suspected condition. If parents choose to terminate the pregnancy, postmortem examination (autopsy) may be offered to confirm or provide further information about the diagnosis. However, postmortem examination may not be offered routinely at all institutions and accepting to perform it is at the discretion of the parents.
If the second-trimester scan results are normal, the pregnancy will follow its natural course and most pregnancies will result in a live birth. All neonates undergo a postnatal physical examination to check the infant's well-being and to screen for congenital anomalies that might necessitate medical intervention or follow-up. However, not all anomalies may present immediately after birth and may only become apparent during the late neonatal period or in childhood. In case of an adverse outcome (i.e. perinatal mortality), parents are often offered the option of a postmortem examination to determine the cause of death.

### First- and second-trimester screening for fetal structural anomalies: two-stage approach



In a two-stage approach, women receive an early anomaly scan (performed at 11 to 14 weeks' gestation) aiming to detect gross abnormalities during the first trimester of pregnancy. This early scan is offered in addition to a routine second-trimester anomaly scan (see Figure 1). If a structural anomaly is detected or suspected at a first-trimester anomaly scan, women are usually referred to a specialist centre for further anatomical evaluation and follow-up testing. If the first-trimester scan is normal, women receive another routine scan during the second trimester of pregnancy. At this point, anomalies may still be identified as during the first trimester of pregnancy not all anomalies are detectable and some may develop at later gestations.[35]

### Rationale

Ultrasound screening policies for fetal structural anomalies vary widely. There is currently no international consensus whether population-based prenatal screening programmes should include a first-trimester anomaly scan, in addition to the routine second-trimester scan traditionally offered. Screening for fetal structural anomalies at 18 to 22 weeks' gestation is thought to allow for an optimal balance between maximising detection rates, while leaving sufficient time for counselling, further investigations, and decision-making prior to 24 weeks' gestation[8,36] - the legal limit for termination of pregnancy in most countries (although exceptions may be made in exceptional circumstances such as lethal conditions).[10] However, the psychological cost following a second-trimester pregnancy termination may be substantial. Long-term follow-up of womenafter termination of pregnancy

has shown that advanced gestational age was associated with higher levels of grief, stress and post-traumatic symptoms, with a substantial number of women developing pathological scores for post-traumatic stress.[37] Long-term psychological morbidity after early termination of pregnancy (before 14 weeks' gestation) was, however, rare.[37] An offer of a first-trimester anomaly scan, as suggested with a two-stage screening approach, intends to detect some gross abnormalities for which termination of pregnancy might be considered early in pregnancy. First-trimester detection of fetal anomalies further offers the advantage of allowing more time for further investigations, counselling and discussions of possible interventions. However, any screening programme has the potential of false positive diagnoses which can cause anxiety throughout the remaining weeks of pregnancy. The potential for false positive diagnoses therefore needs to be evaluated when assessing the performance of potential prenatal screening approaches.[7]

## Objectives

The main objectives of this review are to assess the diagnostic accuracy of first- and second-trimester fetal anomaly screening in low-risk pregnant women, and to compare overall performance of single- and two-stage screening approaches with regards to the number of cases detected before birth, as well as the proportion of false positive diagnoses.

### Secondary objectives

Our secondary objective is to identify potential sources of heterogeneity between studies by investigating the effect of clinical factors (i.e. prior testing, factors related to the conduct of the ultrasound examination, and healthcare setting of the study) and methodological factors (i.e. related to the study design, such as methods used for postnatal ascertainment of cases) on screening performance.

## Methods

### Criteria for considering studies for this review

### Types of studies

We will consider the following study designs for inclusion.

1. Randomised controlled studies in which pregnant women are randomised to different screening strategies for the detection of fetal structural anomalies.
2. Prospective and retrospective cohort studies reporting on the diagnostic accuracy of ultrasound screening for fetal anomalies.
3. Registry-based cohort studies reporting on the number of prenatally and postnatally diagnosed congenital malformations over a period in which a population-based ultrasound screening programme was implemented, and the organisation of this screening programme can be discerned.

Additional inclusion criteria are as follows.

1. The study uses a low-risk or unselected population.
2. A routine screening ultrasound scan (i.e. a scan that is offered to all women, without referral) is performed between 11 + 0 and 13 + 6 and/or 18 + 0 and 23 + 6 weeks' gestation.
3. The reference standard used by the study is described and includes, at a minimum, postnatal examination or postnatal follow-up (or both) of all fetuses that received the index test.
4. Data can be extracted or derived for constructing a 2 x 2 table of the number of true positive, false positive, and true negative test results at the level of each organ system evaluated or each type of anomaly (or both) that occurred in the study population.

Although we strive to extract data to derive the full 2 x 2 table, we anticipate that there will be a paucity of studies from which it will be possible to extract or derive the number of false positive test results. We will therefore include studies from which it is possible to derive data for the full 2 x 2 table, as well as studies from which it is only possible to determine the number of true positives, true negatives, false positives and false negatives. Studies from which it is possible to derive the full 2 x 2 table will be analysed separately.



### Exclusion criteria

1. Studies that evaluate the diagnostic accuracy of targeted ultrasound examination of a specific organ system or subset of anomalies (e.g. targeted fetal echocardiography or dedicated fetal neurosonography)
2. Case-control study designs
3. Literature reviews, case reports, conference abstracts

### Restriction of publication year

Since the initial introduction of ultrasound screening for fetal abnormalities in routine prenatal care, important changes to both the content and conduct of this examination have taken place. An important addition to fetal anatomy assessment has been the introduction of the three vessel view.[38] The introduction of this view to standard ultrasound screening protocols has had a major impact on the detection of congenital heart defects[14] and is now considered a standard screening view for evaluating the fetal heart. Given its impact on the detection of congenital heart defects, and the relatively high prevalence of cardiac anomalies, we have decided to exclude literature preceding the introduction of this view. Therefore, and acknowledging advances in ultrasound technology, we will only consider studies published from 1997 onward. For a review of studies published before 1997, we refer to the HTA by Bricker and colleagues[13] and the NICE guideline[7] on this topic.

### Participants

We will include data from population-based studies, using a low-risk or unselected population, from any country and healthcare setting. A low-risk population will be defined as a study population without known risk factors for fetal abnormalities at the time of inclusion in the study. An unselected population is defined as a study population that includes all pregnant women attending a healthcare centre for routine prenatal care during the study period.

### Index tests

We will consider studies which have evaluated the diagnostic accuracy of first-trimester (11 + 0 to 13 + 6 weeks' gestation) and second-trimester (18 + 0 to 23 + 6 weeks' gestation) ultrasound screening for fetal anomalies, in the setting of single- and two-stage screening approaches.

### Target conditions

The target conditions of interest are fetal structural abnormalities potentially identifiable on ultrasound imaging. Abnormalities for which the detection rate will be determined are listed in Table 1. Minor structural abnormalities, known as soft markers, for chromosomal aneuploidies will not be included in our analyses. Soft markers may be associated with an increased risk for chromosomal abnormalities. However, the soft marker itself as a structural abnormality, has little to no clinical significance and will therefore not be included in this review. We will define first- and second-trimester soft markers for chromosomal abnormalities according to the recommendations published by the American College of Obstetricians and Gynecologists.[39] The anomalies (soft markers) listed below will be excluded from our analyses if reported by studies.

### First-trimester soft markers

1. Increased NT thickness
2. Cystic hygroma

### Second-trimester soft markers

1. Mild or moderate ventriculomegaly (defined as an atrial diameter of less than 15 mm)
2. Choroid plexus cysts
3. Echogenic intracardiac foci
4. Echogenic bowel
5. Mild hydronephrosis or pyelectasis (antero-posterior pelvic diameter measuring 4 mm to 7 mm)
6. Short femur length (measurement at 2.5 centile or below for the gestational age)
7. Thickened nuchal fold (nuchal fold measurement of 6 mm or greater)

**Table 1.** Planned analyses: evaluation of ultrasound accuracy for the detection of fetal structural anomalies

| | Studies | Number of fetuses (cases) | Sensitivity | | | Specificity | | |
|---|---|---|---|---|---|---|---|---|
| | | | Index test 1[a] | Index test 2[b] | Index test 3[c] | Index test 1[a] | Index test 2[b] | Index test 3[c] |
| I. Ultrasound detection of structural abnormalities that are externally visible, that present immediately or shortly after birth with clinically relevant symptoms, or that are considered to be lethal/incompatible with life | | | | | | | | |
| Anomalies included in this analysis are indicated below [*] | | | | | | | | |
| Overall | XX | XXXXX (XXX) | XX | XX | XX | XX | XX | XX |
| Quality of evidence (GRADE): ⊗⊗⊗⊗ | | | | | | | | |
| II. Ultrasound detection of all structural anomalies, categorised according to disease severity | | | | | | | | |
| See Table 2 and Table 3 for anomalies included in each category. | | | | | | | | |
| **Overall** | XX | XXXXX (XXX) | XX | XX | XX | XX | XX | XX |
| *Major anomalies* | *XX* | *XXXXX (XXX)* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Lethal* | *XX* | *XXXXX (XXX)* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Severe* | *XX* | *XXXXX (XXX)* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Moderate* | *XX* | *XXXXX (XXX)* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Minor anomalies* | *XX* | *XXXXX (XXX)* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| Quality of evidence (GRADE): ⊗⊗⊗⊗ | | | | | | | | |
| Lethal: ⊗⊗⊗⊗ \| Severe: ⊗⊗⊗⊗ \| Moderate: ⊗⊗⊗⊗ \| Minor: ⊗⊗⊗⊗ | | | | | | | | |
| III. Ultrasound detection of specific anomalies, categorised per organ system | | | | | | | | |
| [*] Anomalies that are externally visible, present with clinically relevant symptoms shortly after birth, or that are considered to be lethal/incompatible with life | | | | | | | | |
| **CNS** | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Anencephaly ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Spina bifida ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Holoprosencephaly ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Hydrocephalus ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Encephalocele ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| **Respiratory** | XX | XX | XX | XX | XX | XX | XX | XX |
| *CDH** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *CPAM* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| **Cardiac** | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Septal defects †* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Valvular anomaly (biventricular heart) ‡* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Venous return anomalies §* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Aortic arch anomalies ¶* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Conotruncal anomalies *** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *HRHS ††* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |

[a] ***Index test 1:*** *first-trimester ultrasound scan performed as the first part of a two-stage approach.*
[b] ***Index test 2:*** *first- and second-trimester ultrasound scans performed as a two-stage approach combined.*
[c] ***Index test 3:*** *second trimester ultrasound scan performed as part of a single-stage approach.*

*Index test 2 and 3 will be compared, see Table 4.*

| | Studies | Number of fetuses (cases) | Sensitivity | | | Specificity | | |
|---|---|---|---|---|---|---|---|---|
| | | | Index test 1[a] | Index test 2[b] | Index test 3[c] | Index test 1[a] | Index test 2[b] | Index test 3[c] |
| *HLHS* [‡‡] | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Other single ventricle defects* [§§] | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Complex defects with atrial isomerism* [¶¶] | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| Thoracic and abdominal wall | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Gastroschisis ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Omphalocele ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| Gastro-intestinal | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Oesophageal atresia ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Duodenal atresia ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Small bowel obstruction ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| Urinary tract | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Renal agenesis* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Unilateral* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Bilateral ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *MCDK* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Unilateral* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Bilateral ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Other renal dysplasias* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Unilateral* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Bilateral* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Upper urinary tract obstruction* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Lower urinary tract obstruction* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| Genital | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Ambiguous genitalia ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| Musculoskeletal | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Lethal skeletal dysplasia ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Non-lethal skeletal dysplasia ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Limb reduction defects ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Talipes ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| Ear, face and neck | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Orofacial clefts ** | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *Miscellaneous* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| Multisystem anomalies | XX | XXX | XX | XX | XX | XX | XX | XX |
| *Fetal hydrops* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |
| *MCA or syndrome* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* | *XX* |

Abbreviations: US: ultrasound; CNS: central nervous system; CDH: congenital diaphragmatic hernia; CPAM: congenital pulmonary airway malformation; MCDK: multicystic dysplastic kidneys; MCA: multiple congenital anomalies.

* Abnormalities that are externally visible, that present immediately or shortly after birth with clinically relevant symptoms, or that are considered to be lethal/incompatible with life and are included in analysis I.

† Septal defects: e.g. ventricular septal defect(s) (VSD), balanced atrioventricular septal defect (AVSD)*.
‡ Valvular anomalies (biventricular heart): e.g. pulmonary or aortic valve stenosis, Ebstein's anomaly*, tricuspid dysplasia or regurgitation.
§ Venous return anomalies: e.g. total anomalous pulmonary venous return*, partial anomalous pulmonary venous return, isolated persistent left superior vena cava.
¶ Aortic arch anomalies: e.g. aortic coarctation*, hypoplastic or interrupted aortic arch*, double/right aortic arch.
** Conotruncal anomalies: e.g. Tetralogy of Fallot*, double outlet right ventricle*, transposition of the great arteries*, truncus arteriosus*, pulmonary atresia with VSD*
†† HRHS: hypoplastic right heart syndrome*
‡‡ HLHS: hypoplastic left heart syndrome*
§§ Other single ventricle defects: e.g. unbalanced AVSD*.
¶¶ Complex defects with atrial isomerism: e.g. left or right atrial isomerism*, heterotaxy syndromes*.
The cardiac defects listed above are grouped according to their anatomical classification. Examples of the most common types within each anatomical category are provided.

### Reference standards

We will consider several reference standards, depending on the pregnancy outcome: live birth or adverse outcome. For the purpose of this review, we consider any type of fetal or perinatal mortality (i.e. miscarriage, pregnancy termination, intrauterine death, stillbirth, perinatal mortality) as an adverse pregnancy outcome. In live births, the reference standards we will consider include postnatal examination and postnatal follow-up. In cases of an adverse pregnancy outcome, we will consider postmortem examination as the reference standard. However, postmortem examination is not expected to be conducted in all cases of spontaneous pregnancy loss or elective pregnancy termination. Therefore, pregnancies with an adverse outcome are likely not to receive a reference standard, which introduces potential for bias in estimates of test accuracy. Furthermore, pregnancies complicated by structural fetal abnormalities may be more likely to have an adverse outcome (i.e. spontaneous fetal loss, perinatal mortality). In the case of a missing reference standard (i.e. no postmortem examination is performed), both situations will lead to missing false negative test results. We will investigate the impact of likely missing false negative test results by conducting sensitivity analyses where we will inflate the number of false negative results in studies where there is a high number of cases with a missing reference standard, as proposed by Mol 1999[40] and Alldred 2012.[18]



### Search methods for identification of studies

#### Electronic searches

We will search OVID MEDLINE, OVID Embase, the Cochrane Central Register of Controlled Trials (CENTRAL) and Web of Science (Science Citation Index Expanded; Social Sciences Citation Index; Arts & Humanities Citation Index; Emerging Sources Citation Index) using thesaurus terms like MeSH-terms and free text words. The search will consist of two main parts. In part I, we will search for general or prespecified structural abnormalities (listed in Table 1), combined with terms related to ultrasonography, including nuchal translucency and echocardiography, complemented with general terms like scan and prenatal diagnosis, as not all papers may mention the imaging method used. Mild or transient congenital abnormalities that are mainly confined to the third trimester

of pregnancy (e.g. hydronephrosis or pyelectasis) will be combined with a filter for the first and second trimester of pregnancy to restrict these conditions to the pregnancy trimesters of interest as third-trimester screening for fetal anomalies is beyond the scope of this review. In part II, we will only search for terms related to the index tests of interest (e.g. "anomaly scan", "11-14-week scan", etc.) to identify papers that do not mention specific structural anomalies and/or the ultrasound methods and may therefore have been missed by part I. We will combine part I and II with a broad filter for the prenatal period. We will limit our search to studies published after 1997 and we will exclude animal studies, reviews, case reports, and conference abstracts. No further restrictions will be applied. Identified records will be imported into EndNote and duplicate records will be removed. A sample search strategy for MEDLINE is provided in Appendix 1.

### Searching other resources

Reference lists and citing articles of identified relevant papers, including published systematic reviews, will be cross-checked for additional relevant studies, using Web of Science.

### Data collection and analysis

We will apply the methods described in the *Cochrane Handbook for Systematic Reviews of Diagnositc Test Accuracy* (methods.cochrane.org/sdt/handbook-dta-reviews). To ensure consistent and reproducible conduct of the study selection, assessment of methodological quality, and data extraction components of this review, we will develop standardised forms. We will randomly select five publications from the identified studies to pilot our forms and to ensure the criteria are applied consistently.

### Selection of studies

Two review authors will independently screen the titles and abstracts of all studies identified by the search strategy. We will obtain full-text versions of all potentially relevant studies. The full-text versions will be assessed for inclusion by two review authors by using the above-mentioned inclusion and exclusion criteria. If multiple publications report results from the same study cohort, only the results from the most comprehensive and relevant publication will be included. We will resolve discrepancies in selection of studies between review authors by consensus or consultation of a third review author and document reasons for exclusion.

### Data extraction and management

Two review authors will independently extract data from all included studies using the piloted data-extraction form. We will resolve any discrepancies in extracted data between review authors by re-evaluating the data, discussion, or if required consultation of a third author with clinical or methodological expertise depending on the specific query. We will extract the following data from each study.

1. Study characteristics (e.g. reference details allowing for identification of the study, language, study design, study period, healthcare setting, country the study was performed in).
2. Population characteristics and study setting (e.g. sample size, inclusion and exclusion criteria for participation in the study, prior testing).
3. Features of the index test (e.g. primary purpose of the ultrasound scan, type of professional performing the scan, experience level of the operator, time allocated for the ultrasound examination, mode of examination, transducer frequency, use and level of detail of an anatomy checklist (following the classification proposed in Karim 2017[17]), types of cardiac views used, whether fetuses in whom a raised NT was identified received targeted fetal echocardiography, types of anomalies included in the study).
4. Features of the reference standard (e.g. type of professional performing the postnatal examination, length of postnatal follow-up, additional measures taken to identify any missed cases in the study population, postmortem verification of the ultrasound findings in cases of termination of pregnancy for fetal anomaly, postmortem verification of the ultrasound findings in cases of spontaneous pregnancy loss, still birth or perinatal mortality with and without anomalies on ultrasound screening).
5. Data for construction of the 2 x 2 table (number of true positive, false positive, false negative, and true negative test results) or summary statistics from which these data can be derived.
6. Prevalence of structural and chromosomal abnormalities in the study population.

### Assessment of methodological quality



We will use a modified version of the revised QUality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for assessment of methodological quality of included studies[41], tailored to our review question (see Appendix 2). We will omit the QUADAS-2 item assessing the methodological quality of the study according to the time interval between the index test and reference standard as clinically relevant structural fetal abnormalities are expected to be either present or absent. We anticipate that a key methodological issue will be that the reference standard used by studies and its availability will vary depending on the pregnancy outcome. For live-born neonates, postnatal examination is the available reference standard, whereas the available reference standard in cases of adverse pregnancy outcome (any form of fetal or neonatal mortality, including elective pregnancy termination) is postmortem examination. As pregnancies with positive index test results (i.e. an anomaly suspected or detected on ultrasound) are expected to be more likely to result in an adverse pregnancy outcome than pregnancies with negative index test results (normal ultrasound findings), differential verification bias occurs. Further methodological issues are expected to mainly arise from patient flow in studies. Postmortem examination is the reference standard in case of adverse pregnancy outcome. However, acceptance of postmortem examination is reported to be low.[42] As a consequence, we expect many studies to have missing data due to the absence of the reference standard in cases of adverse pregnancy outcome. Again, this might result in the introduction of bias to studies due to the increased likelihood of adverse outcome in pregnancies with positive index test results.

A further potential methodological issue may be the reliability of the reference standard used by studies to diagnose or detect different types of anomalies. We expect that the reference standard used by most studies will involve postnatal examination of the neonate, and follow-up until hospital discharge. Whereas anomalies that are either externally visible, or become symptomatic shortly after birth, are expected to be reliably diagnosed using this reference standard, anomalies that may present later in life will inevitably be missed. Therefore, we will make a distinction in our analyses of overall detection of fetal anomalies between anomalies that are expected to become symptomatic within the first week of life (subgroup 1), and all anomalies taken together (subgroup 2) - see Table 1. Subgroup 2 will be further categorised according to disease severity. We will evaluate the reliability of the reference standard used by studies to diagnose the anomalies included in each of these subgroups in order to provide a more nuanced assessment of the methodological quality of included studies.

### Statistical analysis and data synthesis

Fetal structural anomalies comprise a heterogeneous group of conditions with extensive differences in treatment options, prognoses, and sonographic markers associated with the defect. Important information is lost when only an overall estimate of test accuracy is given in the context of ultrasound screening for fetal anomalies. Therefore, in addition to overall estimates of test accuracy, we will determine the accuracy of ultrasound screening for each organ system evaluated, and for selected malformations occurring in each organ system. Malformations are selected based on their prevalence and medical implications, and can be found in Table 1. If one or more major structural anomalies occurred in the same fetus, they will be coded as "multiple congenital anomaly/syndrome". As mentioned above, we will perform a subgroup analysis for the following main subgroups.

1. Only those anomalies that are expected to present within the first week of life.
2. All anomalies reported by studies.

Subgroup 2 ("all anomalies") will be further categorised according to the severity of the condition, modified after a proposal by The Royal College of Obstetricians and Gynaecologists[11] and Saltvedt 2006.[43] Anomalies included in each severity group are specified in Table 2 and Table 3.
Minor anomalies are defined as conditions or abnormalities associated with minor morbidity or disability. Major anomalies are defined as conditions with serious medical implications that require immediate postnatal evaluation or intervention (or both), and/or are associated with long-term morbidity or disability. Within this category, we recognise the following three subcategories.

1. Lethal anomalies, i.e. conditions that are incompatible with life.
2. Severe anomalies, i.e. conditions that are associated with severe long-term morbidity or severe disability.

3. Anomalies of moderate severity, i.e. conditions associated with short- or long-term morbidity or disability of moderate severity.

**Table 2.** Classification of anomalies according their likely clinical consequences in each organ system (A)

| Classification | | Types of anomalies per organ system | | | | |
|---|---|---|---|---|---|---|
| Main classification | Sub-classification | Central nervous system | Respiratory | Cardiac | Thoracic and abdominal wall | Gastro-intestinal |
| Major | **Lethal anomalies** | Anencephaly | - | - | - | - |
| | **Severe anomalies** | Spina bifida<br>Holoprosencephaly<br>Hydrocephalus<br>Encephalocele | CDH | Critical congenital heart defects [a] | Gastroschisis<br>Omphalocele | Oesophageal atresia<br>Duodenal atresia<br>Small bowel obstruction |
| | **Anomalies of moderate severity** | - | CPAM | Congenital heart defects of moderate severity [b] | - | - |
| Minor | **Minor anomalies** | - | - | Minor congenital heart defects [c] | - | - |

Classification of structural anomalies according their likely clinical consequences. Major anomalies are further categorised into: lethal anomalies (i.e. conditions that are incompatible with life), severe anomalies (i.e. conditions associated with severe immediate or long-term morbidity/severe disability), and anomalies of moderate severity (i.e. conditions associated with short- or long-term morbidity/disability of moderate severity). Minor anomalies are defined as anomalies associated with minor morbidity/disability. Abbreviations: CDH: congenital diaphragmatic hernia; CPAM: congenital pulmonary airway malformation.

[a] Hypoplastic right heart syndrome (any type); hypoplastic left heart syndrome (any type); other single ventricle defects (any type); conotruncal anomalies (any type); total anomalous venous return; Ebstein's anomaly; atrioventricular septal defect (balanced and unbalanced); aortic coarctation; interrupted or hypoplastic aortic arch; complex defects with atrial isomerism (any type).

[b] Isolated pulmonary or aortic valve stenosis of moderate severity; partial abnormal pulmonary venous return.

[c] Other isolated non-severe valvular defects; isolated small ventricular septal defects.

**Table 3.** Classification of anomalies according to their likely clinical consequences in each organ system (B)

| Classification | | Types of anomalies per organ system | | | | |
|---|---|---|---|---|---|---|
| Main classification | Sub-classification | Urinary tract | Genital | Musculoskeletal | Ear, face and neck | Multi-system anomalies |
| Major | Lethal anomalies | Bilateral renal agenesis; Bilateral MCDK | - | Lethal skeletal dysplasia | - | - |
| | Severe anomalies | Other bilateral renal dysplasias<br><br>Severe obstructive uropathies | Ambiguous genitalia | Non-lethal skeletal dysplasia<br><br>Limb reduction defects resulting in severe long-term handicap (e.g. bilateral absence or severe hypoplasia of upper or lower extremities) | - | Fetal hydrops MCA |
| | Anomalies of moderate severity | Unilateral renal agenesis<br><br>Unilateral renal dysplasias (any type)<br><br>Obstructive uropathies of moderate severity | - | Talipes; limb reduction defects resulting in long-term handicap of moderate severity (e.g. unilateral absence or severe hypoplasia of an arm/hand, leg/foot) | Oro-facial clefts | - |
| Minor | Minor anomalies | - | - | - | - | - |

Classification of structural anomalies according their likely clinical consequences. Major anomalies are further categorised into: lethal anomalies (i.e. conditions that are incompatible with life), severe anomalies (i.e. conditions associated with severe immediate or long-term morbidity/severe disability), and anomalies of moderate severity (i.e. conditions associated with short- or long-term morbidity/disability of moderate severity). Minor anomalies are defined as anomalies associated with minor morbidity/disability. Abbreviations: MCDK: multicystic dysplastic kidney disease; MCA: multiple congenital anomalies.

### Statistical analysis and presentation of the data

Individual study results will be presented in coupled forest plots, reporting true positives, false positives, false negatives, true negatives, sensitivity and specificity, where available. From studies that provide insufficient data to construct full 2 x 2 tables (missing false positive rates), we will present data for sensitivity in the forest plots.

Studies that do provide information to construct the full 2 x 2 table will be analysed separately. As the index tests are binary (i.e. the anomaly is seen or not) or have a standardised threshold for positivity (hydrocephalus), we will use the bivariate model to estimate a summary sensitivity and specificity. This summary point and the confidence region around it will be presented in a Summary Receiver Operating Characteristic (SROC) plot. Where possible, we will estimate the summary sensitivity and specificity for each test with respect to detection of congenital anomalies overall, per organ system, and selected malformations as specified in previous sections. If a sufficient number of studies (four or more) without high risk of bias provide information to meta-analyse predictive values (negative or positive, depending on the reported data), we will do so using a random-effects logistic regression (univariate) meta-regression model (only modelling the positive, or the negative predictive value). We will use the total proportion of cases (prevalence) as a covariate.

### Comparison between tests

The main comparison will be between the overall screening performance of single- and two-stage screening approaches (i.e. comparison of the number of cases detected prenatally and false positive diagnoses between cohorts of women that received a single anomaly scan during the second trimester of pregnancy with women who received an anomaly scan during the first and second trimester of pregnancy); see Table 4. To calculate the difference in number of cases detected and false positives for fetal structural abnormalities in a hypothetical cohort of 10,000 pregnant women, we will use an estimate of the prevalence using the pooled prevalence of the pre-specified anomalies we plan to analyse (see Table 1). We will then calculate the difference in sensitivity and specificity between single- and two-stage screening approaches to estimate the difference in the number of cases detected between the approaches, as well as the difference in the number of false positive diagnoses. If more than four studies directly compare the sensitivity and specificity of single- and two-stage screening approaches as defined in the previous sections for at least one anomaly or subgroup, we will perform meta-regression by adding the approach as a covariate to our model (within-study comparison).

### Investigations of heterogeneity

Based on previously published systematic reviews on this topic, we anticipate to encounter substantial heterogeneity between studies. We will initially visually inspect the forest plots and ROC curves to examine the presence of heterogeneity. Sources of heterogeneity will be separated into the following:



1. Study characteristics (e.g. study design, study period, healthcare setting, country the study was performed in).
2. Population characteristics (e.g. inclusion of aneuploid fetuses in the study results, prevalence of congenital anomalies in the study population, whether or not any type of prior testing was received).
3. Features of the index test (e.g. type of professional performing the ultrasound scan, operator experience level, use of an anatomy checklist, time allocated for the ultrasound examination, mode of examination (e.g. transabdominal or transvaginal), transducer frequency, types of cardiac views used (i.e. the four chamber view, left and right ventricular outflow tracts, three vessel view, three vessel and trachea view), types of anomalies included in the study).
4. Features of the reference standard (e.g. type of professional performing the postnatal examination, length of postnatal follow-up, whether or not additional measures were undertaken to identify missed cases).

The data available from included studies will determine the feasibility of performing investigations of heterogeneity for each of the potential sources mentioned. We will perform meta-regression by fitting each potential source of heterogeneity as a covariate in a hierarchical model to identify factors associated

with the diagnostic accuracy of ultrasound screening for fetal abnormalities. We will define for each potential source of heterogeneity whether it is to be fitted as a categorical or continuous variable.

**Table 4.** Planned analyses: comparison between single- and two-stage approaches

| | Studies | Number of fetuses (cases) | Comparison between single- and two-stage approaches (index test 2 and 3) [a] | | | |
|---|---|---|---|---|---|---|
| | | | Sensitivity | Specificity | Cases detected per 10,000 pregnancies | False positive diagnoses per 10,000 pregnancies |
| I. Ultrasound detection of structural abnormalities that are externally visible, that present immediately or shortly after birth with clinically relevant symptoms, or that are considered to be lethal/incompatible with life | | | | | | |
| Anomalies included in this analysis are indicated below [*] | | | | | | |
| Overall | XX | XXXXX (XXX) | +/- XX | +/- XX | +/- XX | XX |
| Quality of evidence (GRADE): ⊗⊗⊗⊗ | | | | | | |
| II. Ultrasound detection of all structural anomalies, categorised according to disease severity | | | | | | |
| See Table 2 and Table 3 for anomalies included in each category. | | | | | | |
| Overall | XX | XXXXX (XXX) | +/- XX | +/- XX | +/- XX | XX |
| *Major anomalies* | *XX* | *XXXXX (XXX)* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Lethal* | *XX* | *XXXXX (XXX)* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Severe* | *XX* | *XXXXX (XXX)* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Moderate* | *XX* | *XXXXX (XXX)* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Minor anomalies* | *XX* | *XXXXX (XXX)* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Quality of evidence (GRADE): ⊗⊗⊗⊗ | | | | | | |
| Lethal: ⊗⊗⊗⊗ \| Severe: ⊗⊗⊗⊗ \| Moderate: ⊗⊗⊗⊗ \| Minor: ⊗⊗⊗⊗ | | | | | | |
| III. Ultrasound detection of specific anomalies, categorised per organ system | | | | | | |
| [*] Anomalies that are externally visible, present with clinically relevant symptoms shortly after birth, or that are considered to be lethal/incompatible with life | | | | | | |
| CNS | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Anencephaly ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Spina bifida ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Holoprosencephaly ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Hydrocephalus ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Encephalocele ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Respiratory | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *CDH** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *CPAM* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Cardiac | XX | XX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Septal defects †* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Valvular anomaly (biventricular heart) ‡* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Venous return anomalies §* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Aortic arch anomalies ¶* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Conotruncal anomalies *** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *HRHS ††* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *HLHS ‡‡* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Other single ventricle defects §§* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |

| | Studies | Number of fetuses (cases) | Comparison between single- and two-stage approaches (index test 2 and 3) [a] | | | |
|---|---|---|---|---|---|---|
| | | | Sensitivity | Specificity | Cases detected per 10,000 pregnancies | False positive diagnoses per 10,000 pregnancies |
| *Complex defects with atrial isomerism* ¶¶ | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Thoracic and abdominal wall | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Gastroschisis ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Omphalocele ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Gastro-intestinal | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Oesophageal atresia ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Duodenal atresia ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Small bowel obstruction ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Urinary tract | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Renal agenesis* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Unilateral* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Bilateral ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *MCDK* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Unilateral* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Bilateral ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Other renal dysplasias* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Unilateral* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Bilateral* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Upper urinary tract obstruction* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Lower urinary tract obstruction* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Genital | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Ambiguous genitalia ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Musculoskeletal | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Lethal skeletal dysplasia ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Non-lethal skeletal dysplasia ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Limb reduction defects ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Talipes ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Ear, face and neck | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Orofacial clefts ** | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *Miscellaneous* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| Multisystem anomalies | XX | XXX | +/- XX | +/- XX | +/- XX | +/- XX |
| *Fetal hydrops* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |
| *MCA or syndrome* | *XX* | *XX* | *+/- XX* | *+/- XX* | *+/- XX* | *+/- XX* |

Abbreviations: US: ultrasound; CNS: central nervous system; CDH: congenital diaphragmatic hernia; CPAM: congenital pulmonary airway malformation; MCDK: multicystic dysplastic kidneys; MCA: multiple congenital anomalies.

* Abnormalities that are externally visible, that present immediately or shortly after birth with clinically relevant symptoms, or that are considered to be lethal/incompatible with life and are included in analysis I. † Septal defects: e.g. ventricular septal defect(s) (VSD), balanced atrioventricular septal defect (AVSD)*.

‡ Valvular anomalies (biventricular heart): e.g. pulmonary or aortic valve stenosis, Ebstein's anomaly*, tricuspid dysplasia or regurgitation.
§ Venous return anomalies: e.g. total anomalous pulmonary venous return*, partial anomalous pulmonary venous return, isolated persistent left superior vena cava.
¶ Aortic arch anomalies: e.g. aortic coarctation*, hypoplastic or interrupted aortic arch*, double/right aortic arch.
** Conotruncal anomalies: e.g. Tetralogy of Fallot*, double outlet right ventricle*, transposition of the great arteries*, truncus arteriosus*, pulmonary atresia with VSD*
†† HRHS: hypoplastic right heart syndrome*
‡‡ HLHS: hypoplastic left heart syndrome*
§§ Other single ventricle defects: e.g. unbalanced AVSD*.
¶¶ Complex defects with atrial isomerism: e.g. left or right atrial isomerism*, heterotaxy syndromes*.
The cardiac defects listed above are grouped according to their anatomical classification. Examples of the most common types within each anatomical category are provided.

### Sensitivity analyses

We will perform sensitivity analyses to examine the effect of the following on the summary estimates of test accuracy.

1. Likely missing false negative test results in pregnancies with an adverse outcome.
2. Registry-based studies.
3. Studies with a small sample size (e.g. less than 10 fetuses with structural anomalies).
4. Studies with "unclear" applicability concerns according to the methodological quality assessment using the modified QUADAS-2 tool.
5. Studies with an overall "high" or "unclear" judgement of risk of bias according to the methodological quality assessment using the modified QUADAS-2 tool.
6. Studies that contain within-study comparisons between single- and two-stage screening strategies.

### Assessment of reporting bias

We will not include a formal assessment of reporting bias in our review due to the existing controversy regarding the assessment and interpretation of reporting bias in test accuracy reviews, especially in the presence of heterogeneity.[44,45]

## References


1. Moorthie S, Blencowe H, Darlison MW, et al. Estimating the birth prevalence and pregnancy outcomes of congenital malformations worldwide. *J Community Genet.* 2018;9(4):387-396.
2. United Nations, Department of Economic and Social Affairs, Population Division. Annual number of births and crude birth rate the 2008 revision of World Fertility Data.www.un.org/en/development/desa/population/publications/dataset/fertility/data.asp (accessed 6 July 2021).
3. Christianson A, Howson CP, Modell B. March of Dimes: global report on birth defects. www.marchofdimes.org Accessed 12 March 2020.
4. Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet.* 2016;388(10063):3027-3035.
5. Boyle B, Addor MC, Arriola L, et al. Estimating Global Burden of Disease due to congenital anomaly: an analysis of European data. *Arch Dis Child Fetal Neonatal Ed.* 2018;103(1):F22-f28.
6. Global Health Observatory (GHO) data: causes of child death. https://apps.who.int/gho/data/node.main.COCD?lang=en (Accessed 6 July 2021).
7. National Collaborating Centre for Ws, Children's H. National Institute for Health and Clinical Excellence: Guidance. In: *Antenatal Care: Routine Care for the Healthy Pregnant Woman.* London: RCOG Press
Copyright © 2008, National Collaborating Centre for Women's and Children's Health.; 2008.
8. Salomon LJ, Alfirevic Z, Berghella V, et al. Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound Obstet Gynecol.* 2011;37(1):116-126.
9. Boyd P, Garne E, Cariati E, et al. EUROCAT (European Surveillance of Congenital Anomalies) Special Report: Prenatal Screening Policies in Europe. In. EUROCAT Central Registry, Newtownabbey: University of Ulster 2010.
10. Lavelanet AF, Schlitt S, Johnson BR, et al. Global Abortion Policies Database: a descriptive analysis of the legal categories of lawful abortion. *BMC International Health and Human Rights.* 2018;18(1):44.
11. The Royal College of Obstetricians and Gynaecologists. Ultrasound screening for fetal abnormalities. Report of the RCOG Working Party. London: RCOG Press, 1997.
12. Working Party. Routine Ultrasound Screening in Pregnancy: Protocol, Standards and Training. Supplement to Ultrasound Screening for Fetal Abnormalities Report of the RCOG Working Party. London: The Royal College of Obstetricians and Gynaecologists, 2000.
13. Bricker L, Garcia J, Henderson J, et al. Ultrasound screening in pregnancy: a systematic review of the clinical effectiveness, cost-effectiveness and women's views. *Health Technol Assess.* 2000;4(16):i-vi, 1-193.
14. Everwijn SMP, van Nisselrooij AEL, Rozendaal L, et al. The effect of the introduction of the three-vessel view on the detection rate of transposition of the great arteries and tetralogy of Fallot. *Prenat Diagn.* 2018;38(12):951-957.
15. Rossi AC, Prefumo F. Accuracy of ultrasonography at 11-14 weeks of gestation for detection of fetal structural anomalies: a systematic review. *Obstet Gynecol.* 2013;122(6):1160-1167.
16. Farraposo S, Montenegro N, Matias A. Evaluation of the role of first-trimester obstetric ultrasound in the detection of major anomalies: a systematic review. *J Perinat Med.* 2014;42(2):141-149.
17. Karim JN, Roberts NW, Salomon LJ, et al. Systematic review of first-trimester ultrasound screening for detection of fetal structural anomalies and factors that affect screening performance. *Ultrasound Obstet Gynecol.* 2017;50(4):429-441.
18. Alldred SK, Deeks JJ, Guo B, et al. Second trimester serum tests for Down's Syndrome screening. *Cochrane Database Syst Rev.* 2012;2012(6):Cd009925.
19. Alldred SK, Takwoingi Y, Guo B, et al. First trimester ultrasound tests alone or in combination with first trimester serum tests for Down's syndrome screening. *Cochrane Database Syst Rev.* 2017;3(3):Cd012600.
20. Alldred SK, Takwoingi Y, Guo B, et al. First and second trimester serum tests with and without first trimester ultrasound tests for Down's syndrome screening. *Cochrane Database Syst Rev.* 2017;3(3):Cd012599.
21. Badeau M, Lindsay C, Blais J, et al. Genomics-based non-invasive prenatal testing for detection of fetal chromosomal aneuploidy in pregnant women. *Cochrane Database Syst Rev.* 2017;11(11):Cd011767.
22. Olde Scholtenhuis MAG, Cohen-Overbeek TE, Offringa M, et al. Audit of prenatal and postnatal diagnosis of isolated open spina bifida in three university hospitals in The Netherlands. *Ultrasound in Obstetrics & Gynecology.* 2003;21(1):48-52.
23. Brent RL. Environmental causes of human congenital malformations: the pediatrician's role in dealing with these complex clinical problems caused by a multiplicity of environmental and genetic factors. *Pediatrics.* 2004;113(4 Suppl):957-968.
24. Donofrio MT, Moon-Grady AJ, Hornberger LK, et al. Diagnosis and treatment of fetal cardiac disease: a scientific statement from the American Heart Association. *Circulation.* 2014;129(21):2183-2242.
25. International Clearinghouse for Birth Defects Surveillance and Research (ICBDSR), National Center on Birth Defects and Developmental Disabilities from the United States Centers for Disease Control and Prevention (CDC), World Health Organization (WHO). *Birth Defects Surveillance: A Manual for Programme Managers.* Geneva: World Health Organization, 2014.

26. Salomon LJ, Alfirevic Z, Bilardo CM, et al. ISUOG practice guidelines: performance of first-trimester fetal ultrasound scan. *Ultrasound Obstet Gynecol.* 2013;41(1):102-113.
27. Audibert F, De Bie I, Johnson JA, et al. No. 348-Joint SOGC-CCMG Guideline: Update on Prenatal Screening for Fetal Aneuploidy, Fetal Anomalies, and Adverse Pregnancy Outcomes. *J Obstet Gynaecol Can.* 2017;39(9):805-817.
28. Van den Hof MC, Smithies M, Nevo O, et al. No. 375-Clinical Practice Guideline on the Use of First Trimester Ultrasound. *J Obstet Gynaecol Can.* 2019;41(3):388-395.
29. Wax J, Minkoff H, Johnson A, et al. Consensus report on the detailed fetal anatomic ultrasound examination: indications, components, and qualifications. *J Ultrasound Med.* 2014;33(2):189-195.
30. Public Health England. NHS Fetal Anomaly Screening Programme Handbook. https://www.gov.uk/government/publications/fetal-anomaly-screening-programme-handbook 2018.
31. PW Soothill, Lo YMD, on behalf of the Royal College of Obstricians and Gynaecologists. Scientific Impact Paper No. 15: Non-invasive Prenatal Testing for Chromosomal Abnormality using Maternal Plasma DNA. The Obstetrician & Gynaecologist 2014;16(2):148-148.
32. Souka AP, Von Kaisenberg CS, Hyett JA, et al. Increased nuchal translucency with normal karyotype. *Am J Obstet Gynecol.* 2005;192(4):1005-1021.
33. Bardi F, Bosschieter P, Verheij J, et al. Is there still a role for nuchal translucency measurement in the changing paradigm of first trimester screening? *Prenat Diagn.* 2020;40(2):197-205.
34. American Institute of Ultrasound in Medicine. AUIM practice guideline for the performance of fetal echocardiography. *J Ultrasound Med.* 2013;32(6):1067-1082.
35. Syngelaki A, Hammami A, Bower S, et al. Diagnosis of fetal non-chromosomal abnormalities on routine ultrasound examination at 11-13 weeks' gestation. *Ultrasound Obstet Gynecol.* 2019;54(4):468-476.
36. Ward P, Soothill P. Fetal anomaly ultrasound scanning: the development of a national programme for England. *The Obstetrician & Gynaecologist.* 2011;13(4):211-217.
37. Korenromp MJ, Christiaens GC, van den Bout J, et al. Long-term psychological consequences of pregnancy termination for fetal abnormality: a cross-sectional study. *Prenat Diagn.* 2005;25(3):253-260.
38. Yoo SJ, Lee YH, Kim ES, et al. Three-vessel view of the fetal upper mediastinum: an easy means of detecting abnormalities of the ventricular outflow tracts and great arteries during obstetric screening. *Ultrasound Obstet Gynecol.* 1997;9(3):173-182.
39. Practice Bulletin No. 163: Screening for Fetal Aneuploidy. *Obstet Gynecol.* 2016;127(5):e123-e137.
40. Mol BW, Lijmer JG, van der Meulen J, et al. Effect of study design on the association between nuchal translucency measurement and Down syndrome. *Obstet Gynecol.* 1999;94(5 Pt 2):864-869.
41. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.
42. Lewis C, Hill M, Arthurs OJ, et al. Factors affecting uptake of postmortem examination in the prenatal, perinatal and paediatric setting. *Bjog.* 2018;125(2):172-181.
43. Saltvedt S, Almström H, Kublickas M, et al. 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.* 2006;113(6):664-674.
44. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol.* 2005;58(9):882-893.
45. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. *In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.* The Cochrane Collaboration, 2010. Available from methods.cochrane.org/sdt/handbook-dta-reviews.

## Supplemental material

***Appendix 1.*** *Example of search strategy*

***Appendix 2.*** *Modified QUADAS-2 tool*