Ultrasound Obstet Gynecol 2024; **64**: 15–27
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.27649.

Detection of non-cardiac fetal abnormalities on ultrasound at 11–14 weeks: systematic review and meta-analysis

J. N. KARIM[1], D. DI MASCIO[2], N. ROBERTS[3] and A. T. PAPAGEORGHIOU[4,5], for the ACCEPTS study#

[1]*Nuffield Department of Women's and Reproductive Health, Jesus College, University of Oxford, Oxford, UK;* [2]*Department of Maternal and Child Health and Urological Sciences, Sapienza University of Rome, Rome, Italy;* [3]*Bodleian Health Care Libraries, University of Oxford, Oxford, UK;* [4]*Oxford Maternal and Perinatal Health Institute, Green Templeton College, University of Oxford, Oxford, UK;* [5]*Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK*

KEYWORDS: congenital anomaly; early pregnancy; first trimester; prenatal diagnosis; screening; sensitivity; specificity

CONTRIBUTION

What are the novel findings of this study?
In this large-scale systematic review and meta-analysis, the overall detection rate for 16 major fetal anomalies on first-trimester ultrasound was 67.33%, with a specificity of 99.99%, although detection rates varied widely for individual anomalies.

What are the clinical implications of this work?
Most common congenital anomalies can be detected at 11–14 weeks' gestation. Parents should be informed that anomalies may be found at this scan and that detection rates vary for different conditions. Practitioners should have an understanding of the conditions amenable to detection and use a standardized approach to optimize detection and minimize false-positive rates.

ABSTRACT

Objectives *To assess the diagnostic accuracy of two-dimensional ultrasound at 11–14 weeks' gestation as a screening test for individual fetal anomalies and to identify factors impacting on screening performance.*

Methods *This was a systematic review and meta-analysis that was developed and registered with PROSPERO (CRD42018111781). MEDLINE, EMBASE, Web of Science Core Collection and the Cochrane Library were searched for studies evaluating the diagnostic accuracy of screening for 16 predefined, non-cardiac, congenital anomalies considered to be of interest to the early anomaly scan. We included prospective and retrospective studies from any healthcare setting conducted in low-risk, mixed-risk and unselected populations. The reference standard was the detection of an anomaly on postnatal or postmortem examination. Data were extracted to populate 2 × 2 tables and a random-effects model was used to determine the diagnostic accuracy of screening for the predefined anomalies (individually and as a composite). Secondary analyses were performed to determine the impact on detection rates of imaging protocol, type of ultrasound modality, publication year and index of sonographer suspicion at the time of scanning.* Post-hoc *secondary analysis was conducted to assess performance among studies published during or after 2010. Risk of bias assessment and quality assessment were undertaken for included studies using the Quality Assessment of Diagnostic Accuracy Studies-2 tool.*

Results *From 5684 citations, 202 papers underwent full-text review, resulting in the inclusion of 52 studies comprising 527 837 fetuses, of which 2399 were affected by one or more of the 16 predefined anomalies. Individual anomalies were not equally amenable to detection on first-trimester ultrasound: a high (> 80%) detection rate was reported for severe conditions, including acrania (98%), gastroschisis (96%), exomphalos (95%) and holoprosencephaly (88%); the detection rate was lower for open spina bifida (69%), lower urinary tract obstruction (66%), lethal skeletal dysplasias (57%) and limb-reduction defects (50%); and the detection rate was below 50% for facial clefts (43%), polydactyly (40%) and congenital diaphragmatic hernia (38%). Conditions with a low (< 30%) detection rate included bilateral renal agenesis (25%), closed spina bifida (21%), isolated cleft*

Correspondence: Prof. A. T. Papageorghiou, Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK (e-mail: aris.papageorghiou@wrh.ox.ac.uk)

#Members of ACCEPTS study group are listed in Appendix S2.

Accepted: 21 February 2024

lip (14%) and talipes (11%). Specificity was > 99% for all anomalies. Secondary analysis showed that detection improved with advancing publication year, and that the use of imaging protocols had a statistically significant impact on screening performance ($P < 0.0001$*).*

Conclusions *The accurate detection of congenital anomalies using first-trimester ultrasound is feasible, although detection rates and false-positive rates depend on the type of anomaly. The use of a standardized protocol allows for diagnostic performance to be maximized, particularly for the detection of spina bifida, facial clefts and limb-reduction defects. Highlighting the types of anomalies amenable to diagnosis and determining factors enhancing screening performance can support the development of first-trimester anomaly screening programs.*

INTRODUCTION

In most developed countries, pregnant women are offered at least two routine ultrasound scans. The first, known as the first-trimester scan, takes place at 11–14 weeks' gestation. Its purposes include confirming fetal viability, establishing an accurate gestational age from the measurement of fetal crown–rump length, identifying multiple pregnancy and determining chorionicity[1–3]. In addition, fetal nuchal translucency (NT) can be measured as part of a combined screening test, along with maternal age, serum free beta human chorionic gonadotropin and pregnancy-associated plasma protein-A, to determine fetal risks for trisomies 21, 18 and 13[1–4]. This test offers approximately 85–90% sensitivity for Down syndrome and other aneuploidies, with a false-positive rate of 5%[5]. The second scan is offered to women at 18–21 weeks, with the main objective of detecting congenital abnormalities[6].

This model of care is being challenged for a number of reasons. First, the distinction between anomaly screening at the first and second scans is artificial, as organogenesis is mostly complete by 10 weeks' gestation[7,8]. In fact, many chromosomally abnormal fetuses have structural malformations, and therefore early anomaly detection is complementary to the objectives of aneuploidy screening[9,10]. Second, technical improvements in ultrasonography now allow for earlier visualization of fetal anatomy[11,12], and studies suggest that many anomalies are detectable at 11–14 weeks[13,14]. Third, the sociopolitical context of fetal anomaly screening has shifted considerably in recent years, as legal decisions affecting reproductive rights have led to renewed conversations around gestational-age restrictions for termination of pregnancy and prompted a re-evaluation of screening practices[15]. Finally, parents favor access to earlier screening[16–18], as it provides early reassurance or diagnosis. An early diagnosis offers clear advantages, such as additional time for genetic testing, multidisciplinary input, discussions around *in-utero* therapy options and informed and balanced decision-making around pregnancy management.

Prior to implementing a screening test, detection rates and false-positive rates should be established[19]. The detection rate for a fetal anomaly depends on the anomaly itself as well as on maternal, fetal, sonographer and equipment-related factors[20,21]. In addition, consideration should be given to ongoing fetal development. For example, cerebellar anomalies are essentially undetectable before 14 weeks' gestation owing to ongoing fetal brain development. Some anomalies are physically so small that they are below the imaging resolution at early gestations. Therefore, it is quite clear that a first-trimester anomaly scan could never be a direct replacement for anomaly screening later in gestation, and any future first-trimester screening program should focus on the type of anomalies amenable to early detection.

We have shown previously that over half of all major cardiac abnormalities are detectable at 11–14 weeks. The objectives of the present study were to determine the screening characteristics of first-trimester ultrasound for non-cardiac fetal anomalies and to understand the impact of logistical screening decisions, such as gestational age at scanning, mode of ultrasound (transabdominal sonography (TAS) and/or transvaginal sonography (TVS)) and use of an anatomical protocol, on screening outcomes in low-risk, unselected or mixed-risk pregnancy populations.

METHODS

The protocol for this systematic review and meta-analysis was developed and registered with PROSPERO (registration number: CRD42018111781) prior to undertaking the database search, selection of studies and data extraction. The review of all studies included in the meta-analysis and the reporting of results were based on the Meta-Analysis of Observational Studies in Epidemiology (MOOSE) checklist, the Synthesizing Evidence from Diagnostic Accuracy Tests (SEDATE) guidance and the Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies (PRISMA–DTA) guidelines[22–25]. The Cochrane Collaboration Systematic Reviews of Diagnostic Test Accuracy handbook was also consulted[26].

The primary outcome of the study was the diagnostic accuracy of two-dimensional ultrasound at 11–14 weeks' gestation for the detection of a predefined selection of 16 major fetal abnormalities, namely: acrania (anencephaly), body-stalk anomaly, holoprosencephaly, encephalocele, severe ventriculomegaly, spina bifida, facial clefts (cleft lip and/or palate), exomphalos (omphalocele), gastroschisis, congenital diaphragmatic hernia, bilateral renal agenesis, lower urinary tract obstruction, lethal skeletal dysplasias, limb-reduction defects, polydactyly and talipes (club foot), using the Centers for Disease Control Birth Surveillance Toolkit classification system where possible[27] (Appendix S1). This selection of anomalies was chosen to

allow for a focused study of a relatively small group of congenital anomalies that could be of interest to a future first-trimester population screening program. It was developed by the members of the ACCEPTS study group (NIHR grant 17/19/10) (Appendix S2) and informed by: (1) anomalies targeted by current second-trimester anomaly screening programs[28]; (2) anomalies considered 'nearly always' or 'potentially' detectable at 11–14 weeks[20]; (3) the prevalence of anomalies[29]; (4) the severity and/or lethality of anomalies; (5) associations with known genetic syndromes; and (6) the possibility of intrauterine therapy following an early diagnosis.

A secondary objective was to determine which factors associated with the screening test might impact on the detection of the 16 predefined anomalies.

Search strategy

A systematic electronic search strategy was designed with the help of a specialist librarian (N.R.) to identify studies evaluating the diagnostic accuracy of two-dimensional ultrasound in the detection of fetal congenital abnormalities at 11–14 weeks' gestation (Appendix S3). The search was developed initially using free-text terms and subject headings related to prenatal screening, early pregnancy and congenital abnormalities in general[21]. In order to increase sensitivity, free-text terms and subject headings for the specific congenital anomalies of interest were also included. The searches were conducted in MEDLINE (OvidSP), EMBASE (OvidSP), Science Citation Index and Conference Proceedings Citation Index-Science (Web of Science Core Collection) and the Cochrane Database of Systematic Reviews and Cochrane Central Register of Controlled Trials (Cochrane Library, Wiley) from 1 January 1998 to 24 August 2022. Articles written in a language other than English, single case reports, case series with fewer than five subjects, conference abstracts, literature reviews, editorials, letters, personal communications and animal studies were excluded within Endnote X9 after full deduplication of references.

Study selection was performed in stages by two independent reviewers (J.N.K. and D.D.). Titles and abstracts of citations obtained from the systematic electronic search were reviewed to identify potentially relevant studies. Full texts were subsequently evaluated to determine their eligibility for inclusion. The reference lists of all eligible studies were screened manually for additional citations not identified by the initial electronic search. Agreement regarding inclusion and exclusion of studies was achieved by consensus between the two reviewers or by consultation with a third reviewer (A.T.P.).

Study selection

Studies reporting on the detection of fetal abnormalities using two-dimensional TVS, TAS or a combination of both approaches in the first trimester of pregnancy were included. Prospective and retrospective observational studies and randomized controlled trials were eligible for inclusion.

Given the context of this work (i.e. population-based screening), only studies in low-risk (as described by the authors), unselected or mixed-risk patient populations were included. Taking this approach of population-based screening, pregnancies that underwent routine screening and were later found to be at high risk (e.g. multiple pregnancy, fetus with increased NT or abnormal karyotype) were included. We excluded studies in which the objective was to enrol only patients at a high *a-priori* risk (as defined by the authors). Examples of such excluded studies are those that enrolled only women with a previously affected pregnancy or a personal or family history of congenital anomaly; those that enrolled only fetuses with increased NT or an abnormality; and those that enrolled only multiple pregnancies.

Every attempt was made to identify publications from the same research groups that shared screened subjects and, in such cases, only the study judged to be the most relevant to our aims or the one with the largest cohort was included.

This review included studies that focused exclusively on the first-trimester ultrasound detection of the predefined abnormalities and those that screened for all types of structural fetal abnormality, as long as the abnormalities of interest were included in the reported cohort and an individual breakdown for each abnormality was reported. Studies that investigated exclusively the use of first-trimester ultrasound for the detection of fetal chromosomal abnormalities were excluded.

According to our previous work[21], the reported gestational age at screening is often ambiguous. Thus, studies conducted at 11–14 weeks may mean 11 + 0 to 13 + 6 weeks, 11 + 0 to 14 + 0 weeks or 11 + 0 to 14 + 6 weeks. In order to ensure a systematic approach, an *a-priori* decision was made to include all examinations completed by the 14th week of post-menstrual gestation (up to 14 + 6 weeks). We included studies based on the intention to screen prior to 14 + 6 weeks, with the understanding that, in real-world clinical practice, a small proportion of scans may fall outside the intended gestational-age window.

The reference standard for determining the accuracy of first-trimester ultrasound assessment was the detection of an abnormality on postnatal or postmortem examination. Studies that did not state an intention to perform such an examination for confirmation were excluded. Studies that aimed to, but did not always, achieve complete follow-up of their patient cohort were still eligible for inclusion in the meta-analysis for pragmatic reasons. Similarly, postmortem examination was not a requirement for the inclusion of individual cases, as this is not always achievable.

Data extraction

From each study, we extracted data from tables or the main text on two independent occasions to reduce the risk

of error in data collection; any discrepancies were resolved by consensus or discussion with the third reviewer. We constructed 2×2 tables in order to calculate the rate of true positives, false positives, true negatives and false negatives for each study and for each of the predefined congenital anomalies.

Additional variables extracted were: first author's name, year of publication, sample size, gestational-age window at the time of screening, population characteristics, study type, patient recruitment details, healthcare setting, index test (i.e. TVS, TAS or both), ultrasound protocol used (including anatomical structures evaluated), time allocated to ultrasound assessment, number of sonographers participating in the study and their level of experience, type of malformations assessed and information regarding postnatal follow-up.

Defining screen positives

Ultrasound screening for fetal abnormalities may result in one of three outcomes: screen negative, the diagnosis of a specific anomaly or the suspicion of an anomaly. The latter two situations represent a 'screen positive' test result, and, for the primary analysis, detection rates were calculated regardless of the index of suspicion. We also recognized that a specific diagnostic 'label' in the first trimester may be modified later in pregnancy, as an anomaly may evolve or be reclassified. These cases could not be fairly considered as either a true positive or a false positive, and were therefore documented separately as 'a change of first-trimester diagnosis'.

Estimation of false-positive rate and specificity

The false-positive rate (and therefore specificity) of first-trimester ultrasound screening is difficult to determine because many fetuses with severe or lethal abnormalities undergo early termination of pregnancy without postmortem confirmation[21]. In order to estimate specificity, reported true-positive results were assumed accurate when these led to termination of pregnancy, even if postmortem confirmation was not available. We feel that this approach is reasonable, as pregnancy termination is unlikely to be offered unless the diagnosis is certain; this is also consistent with previous studies in this area[20,21]. We acknowledge that this practice may lead to underascertainment of the false-positive rate.

Quality assessment of studies

Risk of bias and quality assessment were undertaken for all included studies using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool in four key domains: patient selection, index test, reference standard and flow of patients through the study. This was based on a series of signaling questions developed specifically for this review, with each domain graded as low, high or unclear for risk of bias and applicability (Appendix S4).

Statistical analysis

We performed a meta-analysis of data extracted from eligible studies in two steps. First, for each study, summary statistics with 95% CIs were derived to determine the sensitivity, specificity, positive predictive value and negative predictive value of first-trimester ultrasound for the detection of the individual predefined anomalies. Second, we combined individual study statistics to obtain a pooled summary estimate using a random-effects model. A Haldane–Anscombe correction was used, in which a value of 0.5 was added to cells in any 2×2 table, when required, to avoid a division-by-zero error. Heterogeneity between studies was estimated using the I^2 statistic.

In the meta-analysis for the primary outcome, all patients with any type of screen-positive result (diagnosed or suspected) were included.

Preplanned secondary analyses were then conducted to assess factors that may have an impact on first-trimester screening performance, in subgroups stratified according to the following: (1) ultrasound modality (TAS, TVS or both); (2) publication year of the study; (3) diagnostic certainty (definitive diagnosis *vs* high index of suspicion of an abnormality); and (4) whether a standardized imaging protocol was used for assessment. Subgroup analysis was performed only when there were more than five affected cases in each group. Assessment of the impact of gestational age at the time of first-trimester screening on test sensitivity was planned but not undertaken owing to there being insufficient data at the level of individual fetuses. Based on the analysis of publication year, and in order to better reflect current practice, a *post-hoc* secondary analysis was conducted to assess screening performance for studies published in or after 2010. Statistical analysis was performed using StatsDirect statistical software version 3.3.0. (StatsDirect Ltd, Altrincham, UK); significance was set at $P < 0.05$.

RESULTS

The electronic search yielded 5684 citations following removal of duplicates, of which 202 papers underwent full-text review, resulting in the inclusion of 52 studies (527 837 fetuses) in the meta-analysis (Figure 1, Tables 1 and S1). The studies, which were published between 1999 and 2023, were performed in a variety of healthcare settings, although most ($n = 34$) took place, at least in part, in a university hospital or a tertiary-care center. Twelve studies performed multicenter data collection. Forty-six studies examined patients in the first trimester using a systematic anatomical protocol (Table S2). There were seven study cohorts included in the meta-analysis for which no standardized, routine approach for screening in the first trimester was reported. The methodological quality assessment of the included studies is summarized in Figure S1. Thirty-six studies (450 595 fetuses) were published in 2010 or thereafter.

Screening performance for abnormalities

Across the 52 studies, a total of 527 837 fetuses were screened and 2399 anomalies (belonging to one of the 16 targeted anomalies) were identified (prevalence, 0.45% (95% CI, 0.36–0.57%)). Of these 2399 anomalies, 1498 were detected as a congenital anomaly at the time of first-trimester ultrasound screening (1448 diagnosed, 50 suspected); a further 63 cases were false positives (Table 2). Based on the pooled analysis, first-trimester ultrasound screening had a sensitivity of 67.33% (95% CI, 61.49–72.91%) (Figure 2), a specificity of 99.99% (95% CI, 99.98–100.00%), a positive predictive value of 95.94% (95% CI, 93.12–98.04%) and a negative predictive value of 99.86% (95% CI, 99.81–99.90%). The abnormalities detected in the first trimester represented 70.49% (95% CI, 65.43–75.32%) of all antenatally diagnosed ultrasound abnormalities (i.e. cases detected at 11–14 weeks as a proportion of all cases detected before birth at any gestational age). The respective screening characteristics of the 36 studies published during or after 2010 are also reported in Table 2.

The prevalence of each of the 16 individual anomalies assessed as part of this review are listed in Table 3, and their screening performance is summarized in Table 4 for all studies and Table S3 for studies published during or after 2010. The anomalies were grouped *post-hoc* into those with a detection rate of > 80%, 50–80%, 30–50% or < 30% for ease of interpretation (Table 5). It is notable that exomphalos, particularly when diagnosed in karyotype-negative fetuses, was associated with a higher rate of false-positive findings compared with the other anomalies assessed in the study, representing 75% (47/63) of all false-positive cases reported in the review.

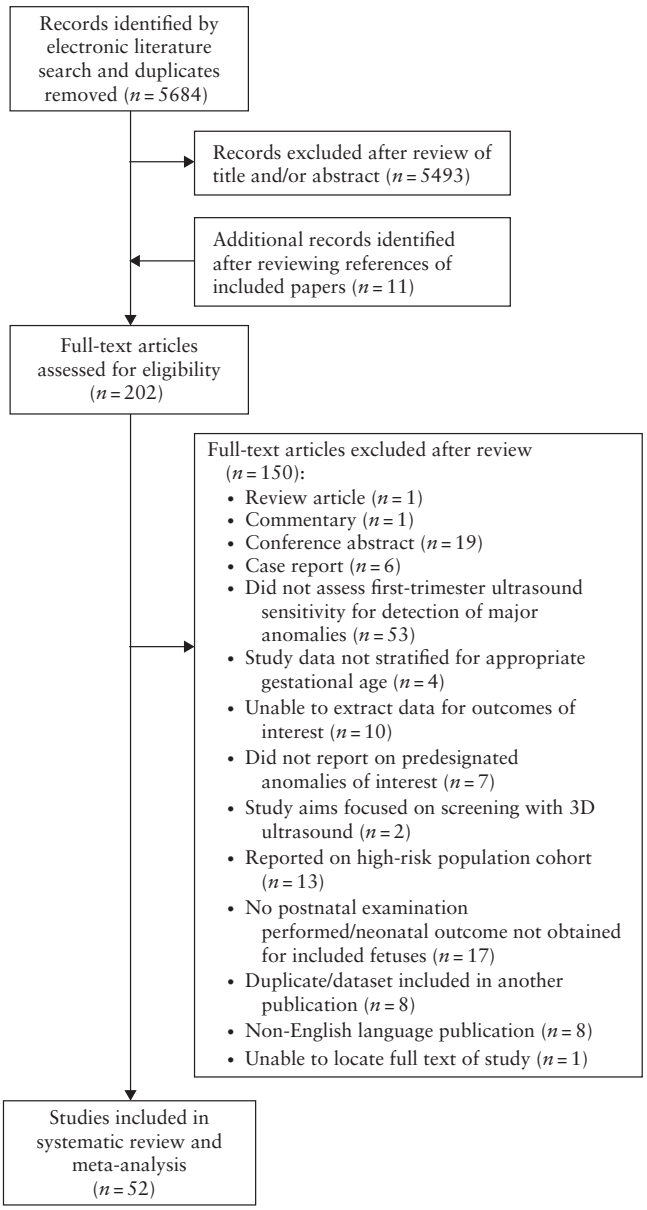

Figure 1 Flowchart summarizing search strategy and selection of studies for inclusion in systematic review and meta-analysis.

Factors affecting screening performance

Ultrasound modality

Most studies used a combination of TAS and TVS (39 study cohorts, 467 736 fetuses), while a few used solely TAS (nine study cohorts, 29 151 fetuses) (Table 1). No studies used a TVS approach alone; five studies did not report on the ultrasound modality. There was no association between ultrasound modality and detection rate ($P = 0.42$) (Table S4).

Publication year

Analysis by year of study publication (in or before 2005, 2006–2010, 2011–2015, in or after 2016) demonstrated improved screening sensitivity with increasing year of publication ($P < 0.0001$) (Figure 2, Table S5).

Diagnostic certainty

Where possible, we differentiated between a definitive diagnosis and a high index of suspicion of an abnormality. The screening performance of first-trimester ultrasound examination according to the degree of diagnostic certainty is shown in Table S6. Of the 1495 anomalies diagnosed in the first trimester, 1448 were true-positive cases and 47 were false-positive cases. Of the 66 anomalies suspected at the time of first-trimester screening, 48 were true-positive cases and 16 were false-positive cases; in two cases, the fetus was affected by an anomaly, but the diagnosis was revised at a later gestational age (change of first-trimester diagnosis). The false-positive rate was higher for suspected compared with diagnosed anomalies (24% *vs* 3%; $P < 0.001$). The positive predictive value, which may be of most relevance to patients, was high when an anomaly was suspected (74%) and when an anomaly was diagnosed (99%).

Table 1 Main characteristics of 52 studies reporting on detection of 16 predefined major fetal anomalies using first-trimester ultrasound (527 837 fetuses)

Study	*Fetuses* (n)	*GA (weeks) or CRL (mm)*	*Aneuploid fetuses included**	*Index test†*	*Anomalies included*
Whitlow (1999)[39]	6634	11 to 14 + 6	Yes (0.7%)	TAS/TVS (20.1%)§	All
Carvalho (2002)[40]	2853	11 to 14	Yes (0.9%)	TAS/TVS	All
Drysdale (2002)[41]	984	12 to 14	Yes	TAS/TVS§	All
Cheng (2003)[42]	3600	10 to 13	Yes (presumed)	NA	Acrania
Taipale (2003)[43]	20 751	11 to 15 + 6	Yes (0.3%)	TAS/TVS (3%)	All
McAuliffe (2005)[44]	325	11 to 13 + 6	No	TAS/TVS (24.6%)	All
Cedergren (2006)[45]	2708	11 to 14	Yes (0.3%)	TAS	All
Saltvedt (2006)[46]	18 053	12 to 14	No	TAS/TVS§	All
Souka (2006)[47]	1148	11 to 14	Yes	TAS/TVS§	All
Srisupundit (2006)[48]	597	11 to 14	Yes	TAS	All
Dane (2007)[49]	1290	11 to 14	Yes	TAS/TVS§	All
Weiner (2007)[50]	1723	11 to 13 + 6	Yes	TAS/TVS (15%)	All
Chen (2008)[51] (control group)	3693	10 to 14 + 6	Yes	TAS/TVS§	All
Chen (2008)[51] (study group)	3949	12 to 14 + 6	Yes	TAS/TVS§	All
Li (2008)[52]	2288	11 to 14	Yes	TAS/TVS (2.0%)§	All
Sepulveda (2008)[53]	5561	45–84	Yes (presumed)	TAS/TVS	CDH
Oztekin (2009)[54]	1085	11 to 14	Yes	TAS/TVS§	All
Abu-Rustum (2010)[55]	1370	11 to 13 + 6	Yes	TAS/TVS§	All
Hildebrand (2010)[56]	6692	11 to 14	Yes (0.2%)	TAS	All
Jakobsen (2011)[57]	9324	11 to 14	Yes	TAS/TVS§	All
Sepulveda (2011)[58]	8936	11 to 13 + 6	Yes	TAS/TVS	Spinal abnormalities (spina bifida and body-stalk anomaly)
Syngelaki (2011)[20]	44 859	11 to 13	No	TAS/TVS (1%)§	All
Adiego (2012)[59]	990	11 to 13 + 6	Yes (presumed)	TAS (TVS in 8%)	Spina bifida
Becker (2012)[60]	6544	11 to 13 + 6	Yes (0.6%)‡	TAS/TVS (23.4%)§	All
Grande (2012)[61]	13 723	11 to 14	No	TAS/TVS	All
Novotna (2012)[62]	9150	11 to 14	Yes	TAS/TVS	All
Pilalis (2012)[63]	3902	11 to 14	Yes	TAS/TVS§	All
Iliescu (2013)[35]	5472	12 to 13 + 6	Yes (0.4%)	TAS/TVS (7.8%)§	All
Sepulveda (2013)[64]	11 068	11 to 13 + 6	Yes	TAS/TVS	All
Wang (2013)[65]	2822	11 to 14	Yes	TAS	All
Natu (2014)[66]	551	11 to 14	Yes	NA	All
Andrew (2015)[67]	4421	11 to 14	Yes	TAS/TVS§	All
Chen (2015)[68]	16 164	11 to 13 + 6	Yes (presumed)	NA	
Colosi (2015)[69]	5924	11 to 13 + 6	Yes (4.7%)	TAS/TVS (1.9%)§	All
Kappou (2015)[70]	2491	11 to 14	Yes (presumed)	TAS/TVS	Spina bifida
Li (2015)[71]	5054	11 to 13 + 6	Yes	TAS/TVS (4%)	Facial clefts
Roman (2015)[72]	23 790	11 to 13 + 6	Yes (presumed)	TAS/TVS	All
Takita (2016)[73]	2028	11 to 13 + 6	Yes (0.6%)	TAS	All
Lakshmy (2017)[74]	2014	50–84	Yes (presumed)	TAS/TVS	Facial clefts
Teegala (2017)[75]	341	11 to 13 + 6	Yes (presumed)	NA	Spina bifida
Vellamkondu (2017)[76]	440	11 to 14	Yes (0.5%)	TAS/TVS	All
Kenkhuis (2018)[34]	5534	11 to 13 + 6	Yes	TAS/TVS§	All
Kose (2018)[77]	1515	45–84	Yes (presumed)	TAS¶	Spina bifida
Vayna (2018)[78]	6114	11 to 14	Yes	TAS/TVS§	All
Zheng (2018)[79]	2982	45–84	Yes	TAS/TVS§	Facial clefts
Chen (2019)[80]	10 294	11 to 13 + 6	Yes	NA	All except spina bifida
Petousis (2019)[33]	3378	11 to 13 + 6	No	TAS/TVS (4%)§	All
Syngelaki (2019)[13]	101 793	11 to 13 + 6	No	TAS/TVS (3%)§	All
Sainz (2020)[81]	504	11 to 14 + 6	Yes	TAS	All
Liao (2021)[14]	59 063	11 to 13 + 6	Yes	TAS/TVS (< 1%)	All**
Liao (2021)[82]	59 063	11 to 13 + 6	Yes	TAS/TVS	Spina bifida
Tiechl (2021)[83]	4949	45–84	Yes (presumed)	TAS	Spina bifida
Li (2023)[84]	7336	45–84	Yes	TAS	Facial clefts

Only first author is given for each study. See Table S1 for complete study characteristics. *In studies that included aneuploid fetuses, proportion of study population confirmed as aneuploid by karyotyping is indicated in parentheses. †In studies in which both transabdominal (TAS) and transvaginal (TVS) ultrasound were used, number in parentheses refers to percentage of study population who received screening with both screening tests. ‡Only known euploid fetuses were included in this meta-analysis, as insufficient data were provided on entire study cohort. §TVS was performed only in situations in which visualization with TAS was suboptimal. ¶TVS was performed only when a structural anomaly was suspected. **Data regarding fetuses affected with spina bifida in this cohort were taken from Liao *et al.*[82]. CDH, congenital diaphragmatic hernia; CRL, crown–rump length; GA, gestational age; NA, not available.

Table 2 Screening performance of first-trimester ultrasound for detection of 16 predefined major fetal anomalies in non-high-risk populations in all studies and in those published during or after 2010

Parameter	*All studies*	*Studies published in or after 2010*
Studies (*n*)	52*	36
Major anomalies (TP + FN) (*n*)	2399	2046
TP (*n*)	1498	1331
FP (*n*)	63	55
Prevalence per 10 000 fetuses (*n*)	45.62 (35.90–56.51)	38.79 (28.56–52.62)
Sensitivity (%)	67.33 (61.49–72.91)	71.98 (65.92–77.67)
Specificity (%)	99.99 (99.98–100.00)	99.99 (99.98–100.00)
PPV (%)	95.94 (93.12–98.04)	96.65 (93.38–98.83)
NPV (%)	99.86 (99.81–99.90)	99.89 (99.84–99.93)
Proportion of all antenatally detected anomalies diagnosed in first trimester (%)	70.49 (65.43–75.32)	76.77 (70.82–82.22)

Data in parentheses are 95% CI. For this analysis, anomalies that were correctly diagnosed or suspected were considered true positives (TP). *52 studies were included but 53 cohorts were assessed independently (two cohorts from Chen *et al.*[51] were analyzed separately). FN, false negative; FP, false positive; PPV, positive predictive value; NPV, negative predictive value.

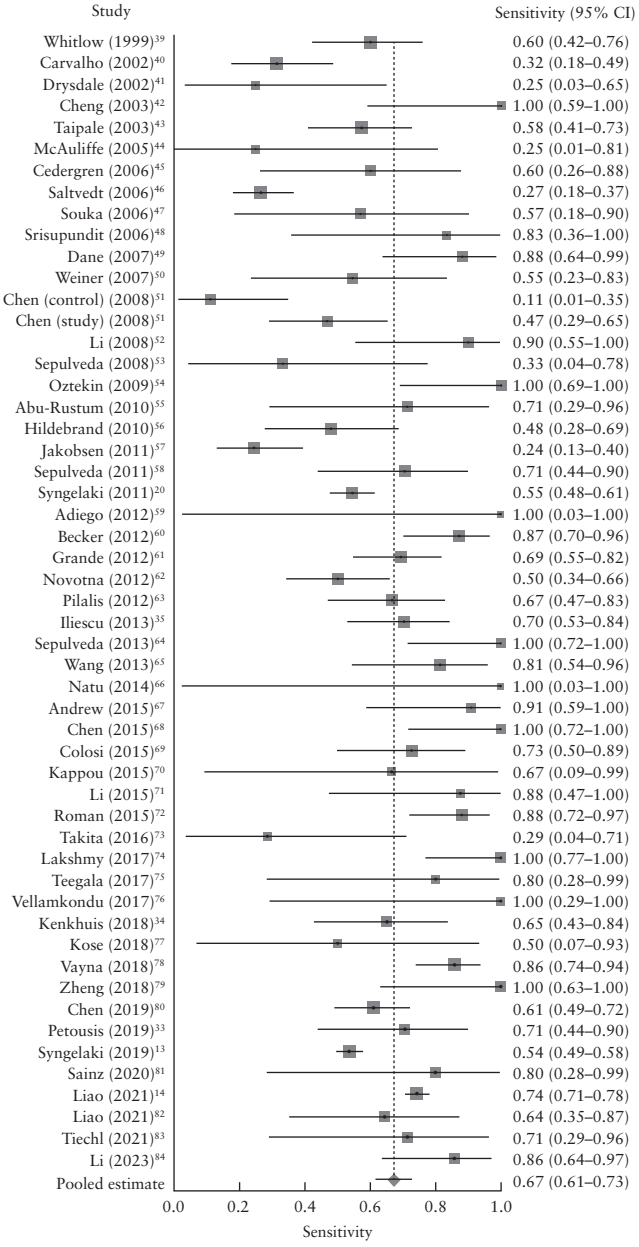

Figure 2 Forest plot of sensitivity of first-trimester ultrasound for detection of 16 predefined major fetal anomalies. Only first author is given for each study. I^2, 84.4% (95% CI, 80.6–87.2%).

Imaging protocol

In a preplanned analysis, detection rates from studies not using a standardized screening approach (six study cohorts) were compared with those using a formal anatomical screening protocol (47 study cohorts (Tables 5 and S7)). The overall detection rate was higher in the latter group (44.55% *vs* 67.55%; $P < 0.0001$).

Assessment of a standardized screening approach in the detection of individual anomalies was challenging given the relatively small number of affected fetuses in the 'no protocol' group (116 anomalies from 25 689 fetuses screened). As analysis was planned only for those anomalies with at least five affected cases in the 'no protocol' and 'with protocol' groups, this was only undertaken for the following anomalies: acrania, gastroschisis, exomphalos, spina bifida (all types), facial clefts (all types), limb-reduction anomalies and talipes (Table 5). A formal anatomical screening protocol was found to have a significant impact on the sensitivity of screening for spina bifida ($P = 0.029$), facial clefts ($P < 0.0001$) and limb-reduction defects ($P = 0.0016$).

DISCUSSION

This meta-analysis of 527 837 screened fetuses focused on the first-trimester detection of 16 anomalies and showed that the majority are identifiable at 11–14 weeks' gestation. First-trimester screening characteristics vary considerably depending on the anomaly under examination, with improved detection rates seen in studies using a standardized anatomical screening protocol. These findings support the feasibility of introducing screening for anomalies during routine first-trimester ultrasound, a practice that would shift the timing of fetal-anomaly diagnosis in most affected women.

Detection of individual anomalies

It was possible to categorize the anomalies based on first-trimester screening sensitivity (Table 5). For acrania, exomphalos, gastroschisis, body-stalk anomaly,

encephalocele and holoprosencephaly, sensitivity exceeded 88%, with detection rates for the first four at approximately 95% or above. For seven anomalies (open spina bifida, lower urinary tract obstruction, lethal skeletal dysplasias, limb-reduction defects, facial clefts, polydactyly, congenital diaphragmatic hernia), a first-trimester diagnosis was achievable in 30–80% of cases, suggesting that these anomalies could reasonably be targeted for detection at 11–14 weeks. The detection of bilateral renal agenesis, severe ventriculomegaly, talipes and closed spina bifida was achievable in < 30% of cases, and detection rates were low even in studies conducted by highly experienced sonographers. These findings are logical and consistent with our understanding of fetal embryology and imaging. The small size of the fetal kidneys makes first-trimester visualization challenging[8], and the absence of amniotic fluid secondary to bilateral renal agenesis is not clinically detectable until after 16 weeks[7]. Severe ventriculomegaly is a progressive condition and, therefore, may not be readily detectable early in gestation[30]. Fetuses with closed spina bifida often present with normal cranial anatomy, thus the cranial sonographic markers reflective of the Arnold–Chiari malformation are unhelpful for this diagnosis[31]. Furthermore, the lack of spinal ossification at this gestational age makes closed spina bifida a challenging anomaly to visualize[7]. Finally, fetal foot orientation is difficult to assess, particularly owing to the non-ossification of the ankle joint at this point in gestation, and the progressive nature of talipes means it may develop later in pregnancy[32]. Our findings suggest that, while first-trimester detection of these anomalies is possible, particularly when the *a-priori* index of suspicion is high or other major anomalies have been diagnosed, the focus of population-level screening for these anomalies should continue to be at later gestational ages. Overall, our findings are consistent with previous classification systems suggesting that anomalies are 'nearly always detectable', 'potentially detectable' or 'virtually undetectable' in the first trimester; our results assess the totality of the evidence and allow a detailed understanding and stratification of the anomalies, particularly those in the 'intermediate' category[20].

Clinical implications

A key objective of this review was to understand the likelihood of false-positive results following first-trimester screening. The majority of constituent studies did not report false-positive rates or were unable to provide a secondary physical confirmation of all detected anomalies. The challenge is that surgical first-trimester pregnancy termination often precludes postmortem examination and, even in prospective studies in which postmortem examination is actively supported, many parents decline this offer[33]. Based on all available data, our best estimate is that the false-positive rate for these conditions is low (3.14%) (Table S6), which is consistent with findings from several individual studies examining this issue[34–36]. It is important to consider which cases should be considered formally to be 'false positives', as anomalies evolve. It is widely accepted that a significant proportion of bowel-only exomphalos and megacystis (≤ 15 mm) cases identified in the first trimester in euploid fetuses will be found to have resolved on imaging at a later gestational age. Within our review, of the 63 reported false-positive cases, 47 (75%) were findings of bowel-only exomphalos in euploid fetuses that subsequently resolved, and when excluding these cases the false-positive rate was 0.13%. Our findings are consistent with those of Iliescu *et al.*[35], who reported a false-positive rate of 3.67% across all types of anomalies, although a more recent study focusing on lethal and severe anomalies suggested a much lower rate of 0.1%[34]. Findings from the Eurofetus Study, which prospectively evaluated second-trimester anomaly screening in 1990–1993, provide additional context[37]. The overall false-positive rate in the Eurofetus Study was

Table 3 Prevalence of 16 predefined major fetal anomalies within study populations included in systematic review and meta-analysis

Anomaly	*Studies screening for anomaly* (n)	*Fetuses screened* (n)	*Affected fetuses* (n)	*Prevalence per 10 000 (95% CI)**
Acrania	40	399 373	314	8.84 (6.69–11.28)
Holoprosencephaly (all types)	38	383 051	131	4.95 (2.91–7.51)
Encephalocele	39	395 773	58	2.38 (1.26–3.87)
Severe ventriculomegaly	37	371 983	23	1.47 (0.11–4.35)
Spina bifida (all types)	46	488 707	220	6.94 (4.97–9.23)
Facial cleft (all types)	41	389 369	392	10.88 (7.66–14.66)
Exomphalos	38	395 773	290	7.45 (5.32–9.94)
Gastroschisis	38	395 773	122	3.23 (2.31–4.32)
Body-stalk anomaly	39	404 709	104	3.97 (1.78–7.02)
Congenital diaphragmatic hernia	38	377 544	86	3.08 (1.94–4.47)
Bilateral renal agenesis	37	371 983	36	1.79 (0.80–3.17)
Limb-reduction defects	37	371 983	128	5.57 (3.25–8.51)
Talipes	37	371 983	308	10.71 (6.70–15.65)
Polydactyly	37	371 983	70	6.81 (2.66–12.86)
Lethal skeletal dysplasias	37	371 983	33	1.49 (0.60–2.76)
Lower urinary tract obstruction	37	371 983	84	2.46 (1.02–4.53)

*Calculated using random-effects model including only studies that screened for anomaly of interest.

9.9%, ranging from 3.1% for spina bifida to 17.1% for abnormalities of the bladder and urethra. In comparison, our findings suggest that the vast improvements in imaging technology seen since the 1990s, combined with higher sonographer skill and increased familiarity with early fetal anatomy, should result in a considerably lower false-positive rate following a first-trimester anomaly scan compared with that seen at the time when second-trimester anomaly screening was first introduced.

In practice, a false-positive result following screening should be distinguished from one that results from a diagnostic test. Studies often do not allow for this distinction; the purpose of early anatomical screening is for stratification into 'high-risk' and 'low-risk' groups. The consequence of a screen-positive result would be referral to a fetal medicine specialist who repeats the ultrasound scan as a diagnostic test that might confirm (true positive) or refute (false positive) the original result. Therefore, it is important that first-trimester ultrasound findings are placed in context for patients, so that they can make the best possible decision for their family with the available information. In the case

Table 4 Screening performance of first-trimester ultrasound for 16 predefined major fetal anomalies

Anomaly	*Study cohorts* (n)	*TP** (n)	*TP + FN* (n)	*FP* (n)	*Sensitivity* (% (95% CI))	*Specificity* (% (95% CI))	*PPV* (% (95% CI))	*NPV* (% (95% CI))
Acrania	34	310	314	0	98.26 (96.57–99.39)	100 (100–100)	98.26 (96.57–99.39)	100 (100–100)
Holoprosencephaly								
All types	22	122	131	0	88.20 (79.75–94.60)	100 (100–100)	96.80 (93.16–99.10)	100 (99.99–100)
Alobar†	7	27	29	0	88.77 (68.93–99.15)	100 (100–100)	94.55 (83.98–99.64)	100 (100–100)
Encephalocele	17	52	58	3	89.94 (81.63–95.95)	100 (100–100)	91.76 (84.01–97.10)	100 (100–100)
Severe ventriculomegaly	4	2	23	0	24.31 (0.29–78.22)	100 (100–100)	52.55 (12.51–90.66)	99.98 (99.97–99.99)
Spina bifida								
All types	38	111	220	5	50.25 (40.25–60.25)	100 (100–100)	89.51 (83.42–94.34)	99.97 (99.96–99.98)
Open†	15	83	137	2	68.79 (50.81–84.25)	100 (100–100)	93.99 (86.65–98.53)	99.98 (99.96–99.99)
Closed†	3	2	11	0	21.15 (3.97–46.97)	100 (100–100)	57.31 (8.54–97.77)	99.96 (99.87–99.99)
Facial cleft								
All types	30	147	392	7	43.44 (28.82–58.66)	100 (99.99–100)	84.22 (73.52–92.55)	99.94 (99.93–99.96)
Cleft lip and palate†	11	16	87	0	42.76 (16.78–71.12)	100 (99.99–100)	84.41 (67.17–95.99)	99.96 (99.92–99.99)
Cleft lip only†	10	5	48	1	13.93 (2.84–31.48)	100 (100–100)	42.78 (19.70–67.66)	99.97 (99.95–99.98)
Cleft palate only†	8	24	52	6	35.34 (9.28–67.46)	99.99 (99.98–100)	60.78 (27.71–89.07)	99.98 (99.98–99.99)
Exomphalos	28	279	290	47	94.73 (91.17–97.41)	99.99 (99.98–100)	91.60 (80.78–98.21)	100 (99.99–100)
Gastroschisis	23	119	122	0	95.64 (91.54–98.43)	100 (100–100)	96.74 (93.03–99.08)	100 (100–100)
Body-stalk anomaly	14	104	104	0	98.00 (94.59–99.76)	100 (100–100)	98.00 (94.59–99.76)	100 (100–100)
CDH	20	32	86	0	38.12 (28.71–48.01)	100 (100–100)	85.26 (73.63–93.94)	99.98 (99.98–99.99)
Bilateral renal agenesis	12	8	36	0	25.30 (13.47–39.37)	100 (100–100)	64.67 (41.51–84.68)	99.99 (99.98–99.99)
Limb-reduction defects	15	75	128	0	50.30 (32.08–68.47)	100 (100–100)	96.10 (90.94–99.16)	99.97 (99.95–99.98)
Talipes	23	27	308	1	11.43 (6.14–18.11)	100 (100–100)	72.21 (55.97–85.90)	99.90 (99.87–99.93)
Polydactyly	11	32	70	0	39.56 (15.01–67.34)	100 (100–100)	78.60 (55.44–94.76)	99.96 (99.94–99.98)
Lethal skeletal dysplasias	10	20	33	0	57.02 (36.78–76.10)	100 (100–100)	88.64 (74.08–97.67)	99.99 (99.99–100)
Lower urinary tract obstruction	13	56	84	0	65.70 (55.66–75.08)	100 (100–100)	93.12 (83.37–98.79)	99.99 (99.98–100)

*Anomalies correctly diagnosed or suspected in first trimester. †As defined by study authors. CDH, congenital diaphragmatic hernia; FN, false negative; FP, false positive; NPV, negative predictive value; PPV, positive predictive value; TP, true positive.

Table 5 Classification of anomalies based on first-trimester screening sensitivity and comparison of detection rates between studies using formal protocol for assessment and those that did not report standardized screening approach

Anomaly	All studies		Sensitivity (%)		P
	Sensitivity (%)	Specificity (%)	No protocol	With protocol	
Pooled screening sensitivity > 80%					
Acrania	98.26	100.00	93.13	97.96	> 0.9999
Body-stalk anomaly	97.99	100.00	NC	NC	—
Gastroschisis	95.64	100.00	91.02	95.95	0.6422
Exomphalos	94.73	99.99	75.75	95.68	0.1911
Encephalocele	89.94	100.00	NC	NC	—
Holoprosencephaly (alobar only)	88.77	100.00	NC	NC	—
Holoprosencephaly (all types)	88.20	100.00	NC	NC	—
Pooled screening sensitivity 50–80%					
Spina bifida (open)	68.79	100.00	NC	NC	—
Lower urinary tract obstruction	65.70	100.00	NC	NC	—
Lethal skeletal dysplasias	57.02	100.00	NC	NC	—
Limb-reduction defects	50.30	100.00	0.00	55.32	0.0016
Spina bifida (all types)	50.25	100.00	25.68	53.52	0.0286
Pooled screening sensitivity 30–50%					
Facial cleft (all types)	43.44	100.00	0.04	49.97	< 0.0001
Cleft lip and palate	42.76	100.00	NC	NC	—
Polydactyly	39.56	100.00	NC	NC	—
Congenital diaphragmatic hernia	38.12	100.00	NC	NC	—
Cleft palate only	35.34	99.99	NC	NC	—
Pooled screening sensitivity < 30%					
Bilateral renal agenesis	25.30	100.00	NC	NC	—
Severe ventriculomegaly	24.31	100.00	NC	NC	—
Spina bifida (closed)	21.15	100.00	NC	NC	—
Cleft lip only	13.93	100.00	NC	NC	—
Talipes	11.43	100.00	8.84	12.49	> 0.9999

NC, not calculated because fewer than five affected cases in 'no protocol' or 'with protocol' group or in any anomaly subgroup.

of bowel-only exomphalos and megacystis diagnosed in the first trimester, healthcare providers should be well informed of the high likelihood of resolution in order to provide parents with appropriate counseling.

Factors impacting screening performance

We have demonstrated previously an association between the sensitivity of early ultrasound for fetal anomalies and the use of a standardized protocol for screening[21]. This has been corroborated by the findings from the present meta-analysis, which demonstrates higher detection rates in studies using a formal anatomical protocol compared with those not using standardized screening ($P < 0.0001$). Meaningful comparisons by anomaly were possible for the most prevalent anomalies, namely acrania, gastroschisis, exomphalos, spina bifida, facial clefts, limb-reduction defects and talipes. These results are important, as they indicate which anomalies would be impacted most by the introduction of a protocolized screening approach. Our findings demonstrate that acrania, exomphalos and gastroschisis are clearly identifiable in the first trimester, even when anomaly detection is not a formal objective. Although a higher proportion of these cases were detected in the first trimester in studies using a protocol, this was not statistically significant, and overall detection rates exceeded 90% in most studies. For anomalies considered 'potentially detectable' in the first trimester, detection rates are influenced by fetal, maternal, sonographer and equipment-related factors. Within this group, an analysis of fetuses affected by spina bifida, those with facial clefts and those with limb-reduction anomalies demonstrated that the sensitivity of first-trimester screening could be improved significantly with the adoption of a standardized screening approach ($P < 0.05$). These anomalies have traditionally been considered challenging to diagnose early in gestation, and it should be acknowledged that the use of a detailed first-trimester protocol to improve detection requires a high level of sonographer skill, the availability of high-resolution ultrasound equipment and appropriate allocation of time. Certainly, the role of sonographic markers in helping to identify at-risk fetuses should be acknowledged and merits further work, for example supporting the detailed examination of the posterior fossa for spina bifida[38] and the evaluation of multiple views of the fetal face for facial clefts.

Most studies ($n = 39$) used a combination of TAS and TVS, making a meaningful comparison against those using only TAS ($n = 9$) difficult (Table S4). Our analysis suggests that there was no significant difference in detection. Studies examining the normal visualization of fetal organs suggest that each modality has advantages and disadvantages, so it is likely that a patient-tailored approach will yield the highest detection rates.

Strengths and limitations

Strengths of our study include the large pooled population of 527 837 fetuses, the use of a prospective, registered protocol and the performance of manual, detailed data extraction. Data collection focused on outcomes from women categorized as low risk, mixed risk or unselected, so were representative of a general obstetric population presenting for routine care. Our review does have some expected limitations. Most of the included studies were from tertiary-care centers, meaning that our findings probably represent the highest level of care and may not reflect routine, daily practice; an element of reporting bias also cannot be excluded. In addition, despite subgroup analysis, considerable heterogeneity between studies was noted, with variations in inclusion and exclusion criteria, age at postnatal follow-up, use of anatomical protocols, experience of sonographers, time allocated to scanning and outcome reporting. Evaluation of the studies using the QUADAS-2 tool found the majority to have a 'high' or 'unclear' risk of bias in relation to the index test, reference test, study flow and timing. An ideal study design would involve blinding the examiner to patient history, prevention of referral bias in tertiary-care centers, postmortem analysis of every terminated fetus and standardized neonatal assessment of internal anomalies in all cases. Such rigorous examination of first-trimester anomaly screening would be very challenging and unlikely to be considered ethical given the high detection rates reported. Nevertheless, it is important to recognize the limitations of the data available within the existing literature.

Conclusions

First-trimester anomaly screening using ultrasound has the potential to detect a number of common congenital anomalies at an earlier gestational age than is the current standard of practice in most settings. The detection and false-positive rates vary depending on the type of anomaly under examination. The development and use of standardized approaches in screening has been shown to optimize detection rates, particularly in the diagnosis of fetuses affected by spina bifida, facial clefts or limb-reduction defects. Understanding the types of anomaly amenable to diagnosis and determining factors that favorably affect screening performance should help in the development of first-trimester anomaly screening programs.

REFERENCES

1. Salomon LJ, Alfirevic Z, Bilardo CM, et al. ISUOG practice guidelines: performance of first-trimester fetal ultrasound scan. *Ultrasound Obstet Gynecol*. 2013;41:102-113.
2. National Collaborating Centre for Women's and Children's Health. *Antenatal Care – Routine Care for the Healthy Pregnant Woman*. RCOG Press; 2008.
3. Bilardo C, Chaoui R, Hyett J, et al. ISUOG Practice Guidelines (updated): performance of 11–14-week ultrasound scan. *Ultrasound Obstet Gynecol*. 2023;61:127-143.
4. Snijders R, Noble P, Sebire N, Souka A, Nicolaides K. UK multicentre project on assessment of risk of trisomy 21 by maternal age and fetal nuchal-translucency thickness at 10–14 weeks of gestation. *Lancet*. 1998;352:343-346.
5. Nicolaides KH. Screening for fetal aneuploidies at 11 to 13 weeks. *Prenat Diagn*. 2011;31:7-15.
6. Ward P, Soothill P. Fetal anomaly ultrasound scanning: the development of a national programme for England. *Obstet Gynaecol*. 2011;13:211-217.
7. Abuhamad AZ, Chaoui R. *First Trimester Ultrasound Diagnosis of Fetal Abnormalities*. Lippincott Williams & Wilkins; 2017.
8. Donnelly JC, Malone FD. Early fetal anatomical sonography. *Best Pract Res Clin Obstet Gynaecol*. 2012;26:561-573.
9. Syngelaki A, Guerra L, Ceccacci I, Efeturk T, Nicolaides KH. Impact of holoprosencephaly, exomphalos, megacystis and increased nuchal translucency on first-trimester screening for chromosomal abnormalities. *Ultrasound Obstet Gynecol*. 2017;50(1):45-48.
10. Kagan KO, Staboulidou I, Cruz J, Wright D, Nicolaides KH. Two-stage first-trimester screening for trisomy 21 by ultrasound assessment and biochemical testing. *Ultrasound Obstet Gynecol*. 2010;36:542-547.
11. Souka AP, Pilalis A, Kavalakis Y, Kosmas Y, Antsaklis P, Antsaklis A. Assessment of fetal anatomy at the 11-14-week ultrasound examination. *Ultrasound Obstet Gynecol*. 2004;24:730-734.
12. Luchi C, Schifano M, Sacchini C, et al. Detailed fetal anatomy assessment in the first trimester at 11, 12 and 13 weeks of gestation. *J Matern Fetal Neonatal Med*. 2012;25:675-678.
13. Syngelaki A, Hammami A, Bower S, Zidere V, Akolekar R, Nicolaides KH. Diagnosis of fetal non-chromosomal abnormalities on routine ultrasound examination at 11–13 weeks' gestation. *Ultrasound Obstet Gynecol*. 2019;54:468-476.
14. Liao Y, Wen H, Ouyang S, et al. Routine first-trimester ultrasound screening using a standardized anatomical protocol. *Am J Obstet Gynecol*. 2021;224:396.e1-396.e15.
15. Scott G. *Nadine Dorries: Bring Down Abortion Time Limit*. The Times; 2022 July 07, 2022;Sect. Health.
16. Maiz N, Burgos J, Barbazan MJ, Recio V, Martinez-Astorquiza T. Maternal attitude towards first trimester screening for fetal abnormalities. *Prenat Diagn*. 2016;36(5):449-455.
17. Bardi F, Bakker M, Kenkhuis MJ, et al. Psychological outcomes, knowledge and preferences of pregnant women on first-trimester screening for fetal structural abnormalities: a prospective cohort study. *PLoS One*. 2021;16:e0245938.
18. Karim J, Craik R, Davidson L, et al. Acceptability of the first trimester anomaly scan amongst parents in the UK. *BJOG*. 2021;128.
19. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. World Health Organization; 1968.
20. Syngelaki A, Chelemen T, Dagklis T, Allan L, Nicolaides KH. Challenges in the diagnosis of fetal non-chromosomal abnormalities at 11–13 weeks. *Prenat Diagn*. 2011;31:90-102.
21. Karim JN, Roberts NW, Salomon LJ, Papageorghiou AT. Systematic review of first-trimester ultrasound screening for detection of fetal structural anomalies and factors that affect screening performance. *Ultrasound Obstet Gynecol*. 2017;50:429-441.
22. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *J Am Med Assoc*. 2000;283:2008-2012.
23. Sotiriadis A, Papatheodorou SI, Martins WP. Synthesizing evidence from diagnostic accuracy TEsts: the SEDATE guideline. *Ultrasound Obstet Gynecol*. 2016;47:386-395.
24. McInnes MD, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA*. 2018;319:388-396.
25. Salameh J-P, Bossuyt PM, McGrath TA, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. *BMJ*. 2020;370:m2632.
26. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 09 0. The Cochrane Collaboration*; 2010:83.
27. Centers for Disease Control and Prevention. Birth Defects Surveillance Toolkit. https://www.cdc.gov/ncbddd/birthdefects/surveillancemanual/index.html2022
28. PHE. *NHS Fetal Anomaly Screening Programme Handbook – Valid from August 2018*. Public Health England; 2018.
29. EUROCAT. Prevalence charts and tables for congenital fetal anomalies (UK) 2013-2019. 2022 [cited January 2021]. https://eu-rd-platform.jrc.ec.europa.eu/eurocat/eurocat-data/prevalence_en: European Commission – European Platform on Rare Disease Registration
30. Melchiorre K, Bhide A, Gika A, Pilu G, Papageorghiou A. Counseling in isolated mild fetal ventriculomegaly. *Ultrasound Obstet Gynecol*. 2009;34:212-224.
31. Ghi T, Pilu G, Falco P, et al. Prenatal diagnosis of open and closed spina bifida. *Ultrasound Obstet Gynecol*. 2006;28:899-903.
32. Abuhamad A. Technical aspects of nuchal translucency measurement. *Semin Perinatol*. 2005;29:376-379.
33. Petousis S, Sotiriadis A, Margioula-Siarkou C, et al. Detection of structural abnormalities in fetuses with normal karyotype at 11–13 weeks using the anatomic examination protocol of the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG). *J Matern Fetal Neonatal Med*. 2020;33:2581-2587.
34. Kenkhuis M, Bakker M, Bardi F, et al. Effectiveness of 12–13-week scan for early diagnosis of fetal congenital anomalies in the cell-free DNA era. *Ultrasound Obstet Gynecol*. 2018;51:463-469.
35. Iliescu D, Tudorache S, Comanescu A, et al. Improved detection rate of structural abnormalities in the first trimester using an extended examination protocol. *Ultrasound Obstet Gynecol*. 2013;42:300-309.
36. D'Ottavio G, Mandruzzato G, Meir YJ, et al. Comparison of first and second trimester screening for fetal anomalies. In: Levi S, Chervenak FA, eds. *Ultrasound Screening for Fetal Anomalies: Is It Worth It?: Screening Revisited after the Eurofetus Data*. Vol 847. Annals of the New York Academy of Sciences; 1998:200-209.

37. Grandjean H, Larroque D, Levi S. The performance of routine ultrasonographic screening of pregnancies in the Eurofetus Study. *Am J Obstet Gynecol*. 1999;181:446-454.
38. Mace P, Mancini J, Gorincour G, Quarello E. Accuracy of qualitative and quantitative cranial ultrasonographic markers in first-trimester screening for open spina bifida and other posterior brain defects: a systematic review and meta-analysis. *BJOG*. 2021;128:354-365.
39. Whitlow BJ, Chatzipapas IK, Lazanakis ML, Kadir RA, Economides DL. The value of sonography in early pregnancy for the detection of fetal abnormalities in an unselected population. *Br J Obstet Gynaecol*. 1999;106:929-936.
40. Carvalho MHB, Brizot ML, Lopes LM, Chiba CH, Miyadahira S, Zugaib M. Detection of fetal structural abnormalities at the 11-14 week ultrasound scan. *Prenat Diagn*. 2002;22:1-4.
41. Drysdale K, Ridley D, Walker K, Higgins B, Dean T. First-trimester pregnancy scanning as a screening tool for high-risk and abnormal pregnancies in a district general hospital setting. *J Obstet Gynaecol*. 2002;22:159-165.
42. Cheng CC, Lee FK, Lin HW, Shih JC, Tsai MS. Diagnosis of fetal acrania during the first trimester nuchal translucency screening for Down syndrome. *Int J Gynaecol Obstet*. 2003;80:139-144.
43. Taipale P, Ammala M, Salonen R, Hiilesmaa V. Learning curve in ultrasonographic screening for selected fetal structural anomalies in early pregnancy. *Obstet Gynecol*. 2003;101:273-278.
44. McAuliffe FM, Fong KW, Toi A, Chitayat D, Keating S, Johnson J-A. Ultrasound detection of fetal anomalies in conjunction with first-trimester nuchal translucency screening: a feasibility study. *Am J Obstet Gynecol*. 2005;193:1260-1265.
45. Cedergren M, Selbing A. Detection of fetal structural abnormalities by an 11-14-week ultrasound dating scan in an unselected Swedish population. *Acta Obstet Gynecol Scand*. 2006;85:912-915.
46. Saltvedt S, Almström H, Kublickas M, Valentin L, Grunewald C. Detection of malformations in chromosomally normal fetuses by routine ultrasound at 12 or 18 weeks of gestation-a randomised controlled trial in 39,572 pregnancies. *BJOG*. 2006;113:664-674.
47. Souka AP, Pilalis A, Kavalakis I, et al. Screening for major structural abnormalities at the 11- to 14-week ultrasound scan. *Am J Obstet Gynecol*. 2006;194:393-396.
48. Srisupundit K, Tongsong T, Sirichotiyakul S, Chanprapaph P. Fetal structural anomaly screening at 11-14 weeks of gestation at Maharaj Nakorn Chiang Mai Hospital. *J Med Assoc Thai*. 2006;89:588-593.
49. Dane B, Dane C, Sivri D, Kiray M, Cetin A, Yayla M. Ultrasound screening for fetal major abnormalities at 11-14 weeks. *Acta Obstet Gynecol Scand*. 2007;86:666-670.
50. Weiner Z, Goldstein I, Bombard A, Applewhite L, Itzkovits-Eldor J. Screening for structural fetal anomalies during the nuchal translucency ultrasound examination. *Am J Obstet Gynecol*. 2007;197:181.e1-5.
51. Chen M, Lee CP, Lam YH, et al. Comparison of nuchal and detailed morphology ultrasound examinations in early pregnancy for fetal structural abnormality screening: a randomized controlled trial. *Ultrasound Obstet Gynecol*. 2008;31:136-146. discussion 146.
52. Li Q, Guan B, Li D. Detection of fetal structural abnormalities by early pregnancy ultrasound in China. *Int J Gynaecol Obstet*. 2008;100:277-278.
53. Sepulveda W, Wong AE, Casasbuenas A, Solari A, Alcalde JL. Congenital diaphragmatic hernia in a first-trimester ultrasound aneuploidy screening program. *Prenat Diagn*. 2008;28:531-534.
54. Oztekin O, Oztekin D, Tinar S, Adibelli Z. Ultrasonographic diagnosis of fetal structural abnormalities in prenatal screening at 11-14 weeks. *Diagn Interv Radiol*. 2009;15:221-225.
55. Abu-Rustum RS, Daou L, Abu-Rustum SE. Role of ultrasonography in early gestation in the diagnosis of congenital heart defects. *J Ultrasound Med*. 2010;29:817-821.
56. Hildebrand E, Selbing A, Blomberg M. Comparison of first and second trimester ultrasound screening for fetal anomalies in the southeast region of Sweden. *Acta Obstet Gynecol Scand*. 2010;89:1412-1419.
57. Jakobsen TR, Sogaard K, Tabor A. Implications of a first trimester Down syndrome screening program on timing of malformation detection. *Acta Obstet Gynecol Scand*. 2011;90:728-736.
58. Sepulveda W, Wong AE, Fauchon DE. Fetal spinal anomalies in a first-trimester sonographic screening program for aneuploidy. *Prenat Diagn*. 2011;31:107-114.
59. Adiego B, Illescas T, Martinez-Ten P, et al. Intracranial translucency at 11–13 weeks of gestation: prospective evaluation and reproducibility of measurements. *Prenat Diagn*. 2012;32(3):259-263.
60. Becker R, Schmitz L, Kilavuz S, Stumm M, Wegner R-D, Bittner U. 'Normal' nuchal translucency: a justification to refrain from detailed scan? Analysis of 6858 cases with special reference to ethical aspects. *Prenat Diagn*. 2012;32: 550-556.
61. Grande M, Arigita M, Borobio V, Jimenez JM, Fernandez S, Borrell A. First-trimester detection of structural abnormalities and the role of aneuploidy markers. *Ultrasound Obstet Gynecol*. 2012;39:157-163.
62. Novotna M, Haslik L, Svabik K, et al. Detection of fetal major structural anomalies at the 11-14 ultrasound scan in an unselected population. *Ceska Gynekol*. 2012;77:330-335.
63. Pilalis A, Basagiannis C, Eleftheriades M, et al. Evaluation of a two-step ultrasound examination protocol for the detection of major fetal structural defects. *J Matern Fetal Neonatal Med*. 2012;25:1814-1817.
64. Sepulveda W, Wong AE. First trimester screening for holoprosencephaly with choroid plexus morphology ('butterfly' sign) and biparietal diameter. *Prenat Diagn*. 2013;33:1233-1237.
65. Wang L, Wu Q-q, Chen Y, Ma Y-q, Yao L, Li M. Ultrasound screening of fetal structural abnormalities by standard ultrasound views during the first trimester. *Chin Med J (Engl)*. 2013;126:986-987.
66. Natu N, Wadnere N, Yadav S, Kumar R. Utility of first trimester anomaly scan in screening of congenital abnormalities in low and high risk pregnancies. *JK Sci*. 2014;16:151.
67. Andrew C, Gopal S, Ramachandran H, Suvika M. First trimester ultrasound: addition of anatomical screening adds value to the examination: a retrospective case series. *J Evol Med Dent Sci*. 2015;4:2690-2700.
68. Chen CP, Wang LK, Chern SR, et al. First-trimester diagnosis of recurrent omphalocele associated with fetal trisomy 18 but without parental mosaicism. *Taiwan J Obstet Gynecol*. 2015;54:194-195.
69. Colosi E, Musone R, Filardi G, Fabbo A. First trimester fetal anatomy study and identification of major anomalies using 10 standardized scans. *J Prenat Med*. 2015;9:24-28.
70. Kappou D, Papastefanou I, Pilalis A, Kavalakis I, Kassanos D, Souka AP. Towards detecting open spina bifida in the first trimester: the examination of the posterior brain. *Fetal Diagn Ther*. 2015;37:294-300.
71. Li W-J, Wang X-Q, Yan R-L, Xiang J-W. Clinical significance of first-trimester screening of the retronasal triangle for identification of primary cleft palate. *Fetal Diagn Ther*. 2015;38(2):135-141.
72. Roman AS, Gupta S, Fox NS, Saltzman D, Klauser CK, Rebarber A. Is MSAFP still a useful test for detecting open neural tube defects and ventral wall defects in the era of first-trimester and early second-trimester fetal anatomical ultrasounds? *Fetal Diagn Ther*. 2015;37:206-210.
73. Takita H, Hasegawa J, Arakaki T, et al. Usefulness of antenatal ultrasound fetal morphological assessments in the first and second trimester: a study at a single Japanese university hospital. *J Med Ultrason (2001)*. 2016;43(1):57-62.
74. Lakshmy RS, Agnees J, Rose N. The upturned superior mesenteric artery sign for first-trimester detection of congenital diaphragmatic hernia and omphalocele. *J Ultrasound Med*. 2017;36(3):583-592.
75. Teegala ML, Vinayak DG. Intracranial translucency as a sonographic marker for detecting open spina bifida at $11–13^{+6}$ weeks scan: our experience. *Indian J Radiol Imaging*. 2017;27:427-431.
76. Vellamkondu A, Vasudeva A, Bhat RG, et al. Risk assessment at 11-14-week antenatal visit: a tertiary referral center experience from South India. *J Obstet Gynaecol India*. 2017;67:421-427.
77. Kose S, Altunyurt S, Keskinoglu P. A prospective study on fetal posterior cranial fossa assessment for early detection of open spina bifida at 11-13 weeks. *Congenit Anom (Kyoto)*. 2018;58(1):4-9.
78. Vayna AM, Veduta A, Duta S, et al. Diagnosis of fetal structural anomalies at 11 to 14 weeks. *J Ultrasound Med*. 2018;37:2063-2073.
79. Zheng MM, Tang HR, Zhang Y, et al. Contribution of the fetal cardiac axis and V-sign angle in first-trimester screening for major cardiac defects. *J Ultrasound Med*. 2018;38:1179-1187.
80. Chen FC-K, Bacovsky A, Entezami M, Henrich W. Nearly half of all severe fetal anomalies can be detected by first-trimester screening in experts' hands. *J Perinat Med*. 2019;47:619-624.
81. Sainz JA, Gutierrez L, García-Mejido J, et al. Early fetal morphological evaluation ($11–13^{+6}$ weeks) accomplished exclusively by transabdominal imaging and following routine midtrimester fetal ultrasound scan recommendations. Since when can it be performed? *J Matern Fetal Neonatal Med*. 2020;33(7):1140-1150.
82. Liao Y, Wen H, Luo G, et al. Fetal open and closed spina bifida on a routine scan at 11 weeks to 13 weeks 6 days. *J Ultrasound Med*. 2021;40:237-247.
83. Tiechl J, Azim SA, Leitner K, et al. Screening for open spina bifida in a routine clinical setting at the first-trimester scan: a prospective multicentre cohort study. *Fetal Diagn Ther*. 2021;48:96-102.
84. Li X, Xiu G, Yan F, et al. First-trimester evaluation of cleft lip and palate by a novel two-dimensional sonographic technique: a prospective study. *Curr Med Imaging*. 2023;19:278-285.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Appendix S1 Data extraction

Appendix S2 Members of the Assessing Clinical and Cost Effectiveness of Prenatal first Trimester anomaly Screening (ACCEPTS) study group

Appendix S3 Electronic search strategy