Systematic review: diagnostic accuracy and outcomes of ultrasound in the first trimester of pregnancy for detection of complications relevant for Austrian population excluding the screening for Down Syndrome [Down's syndrome]

Wilbacher I, Soares-Weiser K, Kleijnen J, Schiller-Frauwirth I, Puig S, Bernardis D, Endel G

CRD summary
This review concluded that good evidence for detecting chromosomal anomalies other than Down's syndrome was found, but test results needed confirmation by karyotyping and availability of karyotyping had to be ensured by the healthcare system. Despite some limitations of the review and included studies, the authors' conclusion seems appropriate.

Authors' objectives
To evaluate the accuracy of ultrasound in the first trimester of pregnancy for diagnosing chromosomal anomalies, chorionicity, risk of preterm birth, gestational diabetes and to determine gestational age.

Searching
MEDLINE, EMBASE, DARE, CINAHL, LILACS and NRR were searched without language restrictions from 1996 to October 2006; search strategies were reported. References of retrieved articles and systematic reviews, internet search engines and contact with experts were used to identify further studies. Update searches were conducted in PubMed in February 2008.

Study selection
Studies that evaluated ultrasound (transvaginal or abdominal) for screening in the first trimester of pregnancy to identify chromosomal anomalies (other than Down's syndrome), placenta chorionicity, risk of preterm birth, gestational diabetes and to determine gestational age compared to a reference standard or repeat screening were eligible for inclusion. Studies that compared ultrasound in the first trimester to ultrasound in the second and/or third trimester were also eligible for inclusion. Diagnostic accuracy studies had to report sufficient data to construct 2x2 tables of test performance. Excluded studies: used predictor tests as the reference standard; evaluated Doppler or echocardiography; were in high-risk populations; or used ultrasound combined with biochemical markers. Karyotyping (with or without pregnancy outcome) or absence of nasal bone were the reference standards for detecting chromosomal anomalies. The cutoff for nuchal translucency ranged from 2.5mm to 4mm and over or in the upper 95% percentile. Most studies were conducted in hospital centres.

Two reviewers selected studies; full papers were not screened independently. Disagreements were resolved by consensus.

Assessment of study quality
Diagnostic accuracy studies were assessed using the 14-criteria QUADAS tool. Trials were assessed in terms of methods of randomisation and allocation concealment, blinding, reporting of eligibility criteria, similarity at baseline, power calculation and use of intention to treat (ITT) analysis. Study quality was assessed by one reviewer and checked by a second.

Data extraction
Data were extracted to produce 2x2 tables of test performance. Sensitivity, specificity, positive and negative likelihood ratios (LR+ and LR-) and diagnostic odds ratio (DOR), with 95% confidence intervals (CI) were calculated. From trials, relative risks and 95% CI for binary outcomes were extracted. Data were extracted by one reviewer and checked by a second.

Methods of synthesis
Statistical heterogeneity was assessed using Cochran Q, I² tests and visual examination of Galbraith plots. Where studies were homogeneous (I²<75%), pooled estimates of sensitivity, specificity and diagnostic odds ratio, with 95% CI, were
calculated using a random-effects model and a summery receiver operating characteristic (SROC) curve was produced (model used not reported) and area under the curve (AUC) was calculated. Where studies were heterogeneous, ranges of diagnostic outcomes were reported and random-effects multivariate regression conducted using the Moses-Shapiro-Littenberg model. A random-effects meta-regression was conducted to investigate potential sources of heterogeneity.

Results of the review
Twenty five studies were stated as meeting the inclusion criteria (number of participants unclear; at least 25 cohort studies and two trials were presented in tables). Of 23 diagnostic accuracy studies assessed: two were retrospective; 18 recruited a representative patient spectrum; 22 used an appropriate reference standard; all avoided progression bias; 18 avoided partial and six differential verification bias; 20 avoided incorporation bias; 21 allowed additional clinical information; and 20 blinded interpreters of the index test. None of the trials reported blinding interpreters of the reference standard. Both trials were considered good quality.

Chromosomal anomalies (including T21) nuchal translucency (10 studies) : Pooled sensitivity 71% (95% CI 67% to 76%) and DOR 86.4 (95% CI 52.1 to 143.3). Specificity ranged from 87% (sensitivity 70%) to 100% (sensitivity 41%), LR+ from 5.54 (LR- 0.35) to 106.5 (LR- 0.59) and LR- from 0.15 (LR+ 35.44) to 0.59 (LR+ 106.5). The AUC was 0.88 (standard error (SE) 0.02).

Chromosomal anomalies (including T21) nasal bone (six studies): Sensitivity ranged from 9% (specificity 100%) to 77% (specificity 99%), specificity from 97% (sensitivity 52%) to 100% (sensitivity 9%), LR+ from 18.81 (LR- 0.49) to 334.1 (LR- 0.23), LR- from 0.23 (LR+ 334.1) to 0.91 (LR+ 20.74) and DOR from 22.7 to 1466.7. The AUC was 0.98 (SE 0.03).

Chromosomal anomalies (excluding T21) nuchal translucency (nine studies): Pooled sensitivity 71% (95% CI 67% to 76%) and DOR 117.3 (95% CI 54.2 to 254.1). Specificity ranged from 87% (sensitivity 50%) to 100% (sensitivity 63%), LR+ from 5.06 (LR- 0.42) to 129.4 (LR- 0.50) and LR- from 0.06 (LR+ 39.15) to 0.50 (LR+ 129.4). The AUC was 0.88 (SE 0.02).

Chromosomal anomalies (excluding T21) nasal bone (six studies): Sensitivity ranged from 30% (specificity 99%) to 88% (specificity 99%), specificity from 97% (sensitivity 33%) to 100% (sensitivity 50%), LR+ from 11.91 (LR- 0.69) to 381.5 (LR- 0.12), LR- from 0.12 (LR+ 381.5) to 0.71 (LR+ 25.05), and DOR from 17.27 to 3235.3. The AUC was 0.99 (SE 0.03).

Results for the multivariate regression for heterogeneous analyses of chromosomal anomalies, results for different ultrasound measurements for detecting chromosomal anomalies (four studies), chorionicity (one study), gestational diabetes (one study) and gestational age (two RCTs) were presented.

Authors' conclusions
Good evidence for detecting chromosomal anomalies other than Down's syndrome was found, but test results needed confirmation by karyotyping and availability of karyotyping has to be ensured by the healthcare system.

CRD commentary
The review addressed a clear research question with appropriate inclusion criteria. Several relevant sources were searched, with attempts to reduce language and publication bias. Diagnostic search filters were used to identify diagnostic accuracy studies, so studies may have been missed. Each stage of the review was conducted in duplicate, which reduced risks of error and bias. Appropriate criteria were used to assess quality and the results were reported for each study. Most of the included diagnostic accuracy studies were subject to differential verification bias and none reported whether results of the index test were known to interpreters of the reference standard. There were several discrepancies between the text and tables and across tables as to the studies included. Appropriate methods of synthesis were used, although it was unclear which model was used to produce the SROC curves and an I^2 of 75% was used as the cutoff that precluded pooling.

Given the data presented, the authors' conclusion seems appropriate.
Implications of the review for practice and research

**Practice:** The authors made several recommendations. Pregnant women must be provided with clear information about the consequences of tests and tests should be optional. Before introduction of first trimester ultrasound, sufficient resources to provide chorion villus sampling within an appropriate time frame were required. Examiners needed adequate training. First trimester ultrasound can not replace second trimester organ screening. Screening for chorionicity using ultrasound was not reasonable. Gestational age should be ascertained prior to 20 weeks gestation. Ultrasound has a role for cervical length assessment in the second and third trimester, but not the first. Additional ultrasound screening should be aimed at avoiding complications during pregnancy and delivery.

**Research:** The authors stated that further research would be necessary to try and reduce the number of false positive results.

**Funding**
Federation of Austrian Social Insurance Institutions.

**Bibliographic details**

**Original Paper URL**
http://www.hauptverband.at/mediaDB/MMDB131571_Report_Ultrasound_Austria.pdf

**Indexing Status**
Subject indexing assigned by CRD

**MeSH**
Austria; Humans; Pregnancy Complications; Pregnancy Trimester, First; Review; Ultrasonography, Prenatal

**AccessionNumber**
12009106279

**Date bibliographic record published**
30/09/2009

**Date abstract record published**
09/02/2011

**Record Status**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.