NIHR | National Institute for Health and Care Research

Health Technology Assessment

Check for updates

Extended Research Article

Clinical and cost-effectiveness of detailed anomaly ultrasound screening in the first trimester: a mixed-methods study

Jehan N Karim,[1] Helen Campbell,[1] Pranav Pandya,[2] Edward C F Wilson,[3] Zarko Alfirevic,[4] Trish Chudleigh,[5] Elizabeth Duff,[6] Jane Fisher,[7] Hilary Goodman,[8] Lisa Hinton,[1] Christos Ioannou,[9] Edmund Juszczak,[10] Louise Linsell,[1] Heather L Longworth,[4] Kypros H Nicolaides,[11] Anne Rhodes,[12] Gordon Smith,[13] Basky Thilaganathan,[14] Jim Thornton,[10] Gillian Yaz,[15] Oliver Rivero-Arias,[1] and Aris T Papageorghiou[1*]

[1]University of Oxford, Oxford, UK
[2]University College London Hospitals NHS Foundation Trust, London, UK
[3]Exeter University, Exeter, UK
[4]Liverpool Women's NHS Foundation Trust, Liverpool, UK
[5]Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
[6]National Childbirth Trust, Cambridge, UK
[7]Antenatal Results and Choices, London, UK
[8]Hampshire Hospital NHS Foundation Trust, Hampshire, UK
[9]Oxford University Hospitals NHS Trust, Oxford, UK
[10]University of Nottingham, Nottingham, UK
[11]King's College Hospital, London, UK
[12]Tiny Tickers, London, UK
[13]University of Cambridge, Cambridge, UK
[14]St George's University Hospitals NHS Foundation Trust, London, UK
[15]SHINE, Peterborough, UK

[*]Corresponding author aris.papageorghiou@wrh.ox.ac.uk

Published May 2025
DOI: 10.3310/NLTP7102

Scientific summary

Clinical and cost-effectiveness of detailed anomaly ultrasound screening in the first trimester: a mixed-methods study

Health Technology Assessment 2025; Vol. 29: No. 22
DOI: 10.3310/NLTP7102

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

In the UK, all pregnant women are currently offered second-trimester ultrasound screening at 18–20^{+6} weeks of gestation for the detection of congenital fetal anomalies. However, many severe and lethal anomalies can be detected earlier and routine first-trimester anomaly screening at 11–14 weeks may be a valuable addition to prenatal care.

Objectives

The objectives of this study were:

1. To assess the diagnostic accuracy of first-trimester ultrasound for major structural anomalies through systematic reviews and meta-analyses of the literature and to understand how this screening should be optimally performed (i.e. anatomical protocol, anomalies to be targeted, gestational age window, ultrasound modality used and referral pathways).
2. To undertake a survey of the current first-trimester screening environment in England.
3. To perform an analysis of UK-based data currently held by the National Congenital Anomaly Disease Registry (NCARDRS) to determine the impact of performing a routine first-trimester anomaly scan on the timing of fetal congenital anomaly diagnosis.
4. To conduct a Delphi consensus procedure for the development of a protocol including technical and logistical aspects of first-trimester anomaly screening, based on expert opinions of healthcare providers from across the UK (sonographers, midwives, obstetricians and fetal medicine specialists).
5. To determine the acceptability of the early anomaly scan among women and their partners.
6. To conduct an economic analysis to estimate the expected costs and outcomes associated with current practice and with prospective first-trimester anomaly screening protocols identified by the work described above.
7. To undertake a value-of-information (VoI) analysis to determine whether there is economic value in undertaking additional future research.
8. To draw together the findings and recommendations from the project, and, if appropriate, outline the design of plausible studies or clinical trials.

Methods

The systematic reviews and meta-analyses of studies were designed to evaluate the diagnostic accuracy of two-dimensional ultrasound for the detection of a pre-selected group of major anomalies at 11–14 weeks' gestation, based on Fetal Anomaly Screening Program (FASP) priorities and the consensus group: anencephaly, holoprosencephaly, encephalocele, body stalk anomaly, ectopia cordis, exomphalos, gastroschisis, lower urinary tract obstruction (LUTO) and major cardiac anomalies. The protocols for the reviews were developed and registered with the International Prospective Register of Systematic Reviews prior to undertaking the search, selection of studies and data extraction (PROSPERO, CRD42018111781 and CRD42018112434). A systematic electronic search strategy was designed with the help of a specialist librarian using free-text terms and subject headings related to prenatal screening, early pregnancy and congenital abnormalities and conducted using four databases (MEDLINE, EMBASE, Web of Science Core Collection and Cochrane Library) for studies published between January 1998 and July 2020. Prospective and retrospective studies evaluating pregnancies of low, mixed or uncertain a priori risk and in any healthcare setting were eligible for inclusion. We excluded studies only evaluating high-risk pregnancies. The reference standard used was the detection of a major abnormality on postnatal or post-mortem examination. Data were extracted from the included studies to populate 2 × 2 tables. Meta-analysis was performed using a random-effects model to determine the performance of first-trimester ultrasound for the detection of the individual pre-selected congenital anomalies, for major cardiac abnormalities overall and for the major non-cardiac anomalies overall (n = 7). Pre-planned secondary

analyses were conducted to assess factors that may impact screening performance, including the imaging protocol used for assessment, ultrasound modality, year of publication, and the index of sonographer suspicion at the time of the scan. Risk of bias and quality assessment were undertaken for all included studies using the quality assessment of diagnostic accuracy studies (QUADAS-2) tool.

The nationwide survey of NHS practice was developed and undertaken in collaboration with the FASP. Thirty-six questions covered domains including current first-trimester ultrasound protocols; local policies regarding screening logistics (e.g. time allocated for scan, mode of scan, equipment) and referral pathways; inclusions of a routine early fetal anomaly scan and resource availability. After validation and piloting, the survey was distributed electronically in January 2019 to all NHS maternity trusts in England (n = 132). Anonymised data were analysed using descriptive statistics for the group of responding trusts; survey responses from trusts in different regions [as defined by Public Health England (PHE) at that time] were compared using chi-squared tests.

Data obtained from the nationwide survey of NHS practice regarding the first-trimester anomaly screening protocols of different NHS trusts were linked to retrospective data held by the NCARDRS from pregnancies with estimated delivery dates between April 2017 and 2019. Ethics approval for this work was obtained after full review by the North West – Preston Research Ethics committee (21/NW/0173) in March 2021 and by the National Disease Registry Project Review Panel on behalf of PHE. Data from NHS Hospital trusts who responded to the nationwide survey were aggregated into one of four groups based on the reported type of first-trimester anomaly screening protocol used routinely: (1) no formal assessment; (2) basic anatomical assessment (routine evaluation of fetal head, limbs and/or cord insertion only); (3) advanced anatomical protocol (basic + either stomach and/or bladder); (4) extended anatomical protocol (advanced + fetal heart, spine and/or face). The primary objective of the study was to determine the proportion of anomalies (a pre-designated group) which are currently identified prior to 16 weeks in England and to compare the early detection rates of these anomalies based on the first-trimester screening protocol used (Group a vs. Group b vs. Group c vs. Group d). The pre-designated anomalies of interest were based on current FASP second-trimester guidance and on several anomalies of interest in the first trimester which included anencephaly, alobar holoprosencephaly, encephalocele, exomphalos, gastroschisis, spina bifida, facial clefts, congenital diaphragmatic hernia, bilateral renal agenesis, megacystis, lethal skeletal dysplasias, limb reduction defects, hypoplastic left heart syndrome (HLHS), atrioventricular septal defect (AVSD), tetralogy of Fallot (TOF) and transposition of great arteries. Pre-specified subanalysis of each type of anomaly by ultrasound protocol was also assessed. Analysis of data at individual trust level was not undertaken.

The Delphi consensus procedure took place entirely online over two rounds using RedCap software (Vanderbilt, Nashville, TN, USA). The study was open to all UK healthcare professionals with an interest in this area of research, with invitations to participate circulated to a list of UK-based sonographers, midwives and doctors with known interests in this area, and to the membership of the British Medical Ultrasound Society and the British Maternal Fetal Medicine Society. All data collected from participants were kept confidential, analysed anonymously and in aggregate form. A literature search conducted from 1991 to 2021 identified (1) all published first-trimester ultrasound protocols evaluating fetal anatomy; (2) a list of anomalies detectable at 11–14 weeks; and (3) relevant screening factors; this formed the basis for round one of the Delphi questionnaire. Participants were asked to identify those fetal anomalies and anatomical views which should be routinely evaluated in the first trimester, and determine logistical aspects. Items receiving ≥ 80% support and < 60% support were included and excluded, respectively, from the protocol. In round two, results were fed back to the participants for confirmation, and items receiving between 60% and 80% support were reconsidered. Subgroup analysis was performed to determine whether responses differed by stakeholder group.

The ACceptability of the first trimester Anomaly Scan (ACAS) Study was a multicentre prospective, questionnaire-based study designed to explore parental views towards routine anomaly screening at 11–14 weeks in the UK. It included two distinct study cohorts. In Cohort A, parents attending routine antenatal ultrasound at 1 of 10 participating NHS hospitals in England and Wales were eligible for recruitment. In Cohort B, parents with a previous pregnancy or child with a congenital anomaly were invited to participate via two national charities: Antenatal Results and Choices and Spina Bifida, Hydrocephalus, Information, Networking, Equality Charity. All participants received a briefing guide explaining the potential benefits and risks of an 11–14 week anatomy assessment and were asked to complete a validated, structured questionnaire on their views regarding screening for anomalies.

For the health economic evaluation, a detailed decision-analytic model was developed to simulate the impact upon healthcare costs and maternal quality-adjusted life-years (QALYs) of a policy to add a first-trimester anomaly scan to the current antenatal screening pathway. Assessments of the impact of the screening policy upon pregnancy outcomes and infant costs and QALYs were also made and are reported separately. Costs included additional time for consent and scanning, sonographer training, as well as additional fetal medicine and echocardiographic scans, and other follow-up investigations offered following an initial screen-positive scan. The implications for maternal quality of life of screening outcomes, further investigations, pregnancy continuation decisions, and fetal losses during the first and second trimesters were also modelled. The model was run for a period of 20 years using an NHS perspective, and populated using data from the project's systematic reviews and surveys, administrative databases, the National Schedule of NHS Costs (2019–20) and the published literature. Parameters were entered using distributions to facilitate probabilistic sensitivity analysis. VoI analysis, conducted on the cost-effectiveness results generated using maternal healthcare costs and QALYs, was used to identify uncertainty present in groups of key model parameters and whether investments in further research are needed to reduce such uncertainty before a policy decision can be made about the implementation of first-trimester anomaly screening.

Results

Based on systematic review of low-risk and unselected pregnancies (416,877 fetuses in 40 studies), for the group of major anomalies prioritised by FASP and the consensus procedure, a first-trimester anomaly scan will detect 93.29% [95% confidence interval (CI) 90.37% to 95.71%] of anomalies with a specificity of 99.99% (95% CI 99.98% to 99.99%) and a positive predictive value (PPV) of 96.54% (95% CI 93.27 to 98.76). False-positive (FP) rates are low, and this is consistent with findings from several individual studies examining this issue. Within our review, there were 49 reported FP cases identified, of which 47 were described as findings of bowel-only exomphalos on first-trimester ultrasound in euploid fetuses which were labelled as having subsequently 'spontaneously resolved'. It should be noted that FP screening will result in additional referrals for fetal medicine assessment, and this has been taken into account within the health economic analysis. For major cardiac anomalies (306,872 fetuses, 45 studies), a first-trimester anomaly scan will detect 55.80% (95% CI 45.87% to 65.50%) of anomalies with a specificity of 99.98% (95% CI 99.97% to 99.99%) and a PPV of 94.85% (95% CI 91.63% to 97.32%). Individually, the first-trimester detection rates for seven of the non-cardiac anomalies in question (acrania, exomphalos, gastroschisis, body stalk anomaly, holoprosencephaly and ectopia cordis) exceed 85% of cases, with fetuses affected by LUTO identified in 65% of cases. We compared studies using a formal anatomical protocol to those not doing so. This showed no statistically significant differences in the detection rates for these eight anomalies combined, nor in the detection of the anomalies individually, with the exception of screening for holoprosencephaly. For major cardiac anomalies, we found strong evidence that the imaging protocol used for examination impacts screening performance ($p < 0.0001$), with a significantly higher detection rate observed in studies using at least one outflow tract view or colour flow Doppler imaging (both $p < 0.0001$). Different types of cardiac anomalies were not equally amenable to detection, though first-trimester detection rates exceeded 70% for the following anomalies: complex cardiac defects, left and right hypoplastic syndromes, arterio-ventricular septal defects, tricuspid atresia, truncus arteriosus and heterotaxy syndromes.

Despite an absence of national recommendations, approximately 75% of units in the UK already perform some form of early anomaly screening, and the majority of trusts do this within the current time allocation of 25–30 minutes. However, significant variations in practice were seen with 64% of trusts using a locally developed anatomical protocol of varying detail, 36% offering in-house sonographer training and 24% giving patients local written pre-scan information specific to first-trimester anomaly screening. There were important differences seen between the services offered across different geographical regions of the UK, resulting in inequity of care.

Data from NCARDRS suggest that NHS hospitals undertaking first-trimester anomaly screening provide significantly more patients with an early diagnosis (before 16 weeks of gestation). The highest detection rates were seen in those centres performing detailed first-trimester ultrasound scans routinely, using formalised protocols (Group d, 40%), but a sizeable proportion of anomalies are also being diagnosed at early gestations in units where no first-trimester anatomy assessment is formally declared (Group a, 28%). A significant association was demonstrated between the sensitivity of early ultrasound at a population level and the use of an anatomical protocol for screening. This suggests that higher detection rates for the pre-designated group of major anomalies are achieved in those centres with the most detailed protocols for screening ($p < 0.001$).

Based on a Delphi consensus procedure, an anatomical protocol for first-trimester screening was developed with the expert opinion of 172 UK healthcare providers recommending that early anomaly screening should be performed at 12–14 weeks' gestation, primarily using transabdominal ultrasound. At a minimum, this screening should target the diagnosis of eight major anomalies: anencephaly, body stalk anomaly, ectopia cordis, encephalocele, exomphalos, holoprosencephaly, gastroschisis and LUTO.

The ACAS Study included participation from 1374 parents (1199 in Cohort A and 174 in Cohort B.) The vast majority of parents felt that first-trimester anomaly screening would be beneficial and would opt for an 11- to 14-week anomaly scan in a future pregnancy (A: 91%, B: 95%). This includes couples who would opt against screening for chromosomal abnormalities and those who would not consider termination of pregnancy. Of note, many parents wish to be informed of a suspected anomaly in the first trimester, even if it cannot be confirmed until a later gestation (A: 74%, B: 82%).

Health economic analysis showed that first-trimester anomaly screening was associated with a small, per woman, mean cost increase of £11 (95% CI £1 to £29) on account of increased scanning times. A mean maternal QALY gain of 0.002065 (95% CI 0.000565 to 0.00358) was driven largely by the temporary reassurance provided by a negative first-trimester anomaly scan to around 90% of all women. The incremental cost-effectiveness ratio was £5270 per maternal QALY and the likelihood of first-trimester anomaly screening being cost-effective at a willingness to pay of £20,000 per QALY was 95%. The model predicted an increase in first-trimester terminations, and reductions in second-trimester terminations and live births of infants with anomalies. These changes led to reductions in infant healthcare costs and QALYs. Maternal and infant costs and QALYs were not aggregated for methodological and ethical reasons, and because of a general lack of guidance around how to interpret the overall implications of such antenatal screening programmes.

The VoI analyses indicated the expected value of perfect information (i.e. the value of removing all uncertainty across all 175 model parameters) for England over a period of 20 years to be £3,461,151. Parameters for the extra costs of anomaly screening (encompassing sonographer training and additional screening time) and the screening performance for the eight anomalies (sensitivities and FPs) accounted for most uncertainty but the value likely to be realised from reducing the uncertainty around these parameters was considered to be lower than the costs of the research needed to achieve this.

Conclusions

Given a framework of standardisation and training, first-trimester ultrasound screening for fetal anomalies is clinically effective and acceptable to parents. Analysis modelling maternal healthcare costs and QALYs indicate that the addition of first-trimester anomaly screening to the current antenatal screening pathway is likely to represent a cost-effective use of resources. Fewer live births of babies with anomalies are predicted and this raises complex and ethically sensitive issues previously documented by analysts evaluating antenatal screening programmes for fetal anomalies. VoI analysis on maternal costs and QALYs suggests that decision uncertainty is low, and that investing in new research to further reduce this uncertainty would not be a cost-effective use of resources. Overall, our report suggests that first-trimester ultrasound screening for fetal anomalies is clinically effective and cost-effective, and that further prospective studies would not constitute an efficient investment.

Study registration

This study is registered as PROSPERO CRD42018111781 and CRD42018112434.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 17/19/10) and is published in full in *Health Technology Assessment*; Vol. 29, No. 22. See the NIHR Funding and Awards website for further award information.

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 3.5

A list of Journals Library editors can be found on the NIHR Journals Library website

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.5 and is ranked 30th (out of 174 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal

Manuscripts are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This article

The research reported in this issue of the journal was funded by the HTA programme as award number 17/19/10. The contractual start date was in August 2018. The draft manuscript began editorial review in January 2023 and was accepted for publication in November 2023. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).