Study Protocol

Title: Artificial Intelligence algorithm for the screening of abnormal fetal brain findings at first trimester ultrasound scan

Acronym: **AIRFRAME 1 GR-2021-12374064 (A**rtificial **I**ntelligence for **R**ecognition of **F**etal b**R**ain **A**no**M**ali**E**s First Trimester)

NCT CE 5526

Date of Approval 16th of February 2023

Principal Investigator (PI)

Dott.ssa Alessandra Familiari, Fondazione Policlinico A. Gemelli, IRCCS, Rome, Italy

Co-Principal Investigator (Co-PI)

Dott.ssa Silvia Salvi, Fondazione Policlinico A. Gemelli, IRCCS, Rome, Italy

Collaborating Centers and PIs

- Prof. Andrea dall'Asta, Parma
- Dr.ssa Tiziana Fanelli, Bari

Funding: Ministero della Salute

Background

Ultrasound fetal examination is an essential and well-established part of the antenatal care. Routine first trimester ultrasound scan offers the opportunity to evaluate the viability of pregnancy, determine the gestational age by measuring the fetal crown-rump length (CRL), assess the risk of fetal aneuploidies by measuring the nuchal translucency thickness (NT), and detect major fetal abnormalities (1). The latter is mainly due to ongoing technological advances, that have allowed the improvement in the resolution of ultrasound imaging up to a point, at which is possible to assess early fetal brain anatomy and detect malformations (1). Early recognition of fetal malformations allows fetal medicine experts to properly counsel the couple about the possible management options for the pregnancy and to plan individualized antenatal and postnatal care. However, in respect of the diagnosis of congenital anomalies first trimester ultrasound has several limitations, such as lower detection rate compared to second trimester ultrasound scan, higher interobserver variability due to the level of experience of each operator, and the later development of some anatomical structures and pathologies (e.g. corpus callosum defects, gastrointestinal obstructions).

Congenital brain and spine malformations are among the most common congenital anomalies. The prevalence of all congenital nervous system anomalies is approximately 27.5 cases in 10.000 births while the prevalence of spina bifida is approximately 4.41/10.000 births (2). Although limited data exist on the incidence of Dandy-Walker malformation (DWM) and cystic anomalies of the posterior fossa, this is estimated to be around 1/11.574 births (3). The antenatal detection of such anomalies plays an essential role because of the possibility to offer genetic investigations or in selected cases fetal surgery (such as open spinal dysraphism), which has been shown to improve the perinatal outcomes. Interestingly both open spina bifida and Dandy Walker malformation are widely known to be characterized by abnormal sonographic findings on the midsagittal view of the fetal brain at the level of the posterior fossa. On this regard, offering routine first trimester ultrasound assessment of the posterior fossa would be valuable to identify early these abnormalities (4).

The routine ultrasound evaluation of the fetal brain at 1st trimester involves some well-defined scanning planes as recommended by the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) (1). Among these, the midsagittal plane of the fetal brain is of paramount importance for the screening chromosomal abnormalities by means of the NT. Another feature of this plane is the visualization, on the posterior region of the fetal brain, of three hypoechoic areas representing, from anterior to posterior, the brain stem (BS), the fourth ventricle (also called intracranial translucency or IT), and the cisterna magna (CM) (5). These three spaces appear clearly delineated and vary in sizes, being the brain steam the widest one, and the CM the narrowest one and the fourth ventricle or IT always visible with a median antero-posterior diameter between 1.5 to 2.5 mm (5), as shown in Figure 1. Moreover, changes in their appearance or size have been reported to be abnormal at 11-13 weeks in fetuses with open spinal dysraphism or cystic anomalies of the posterior fossa. In fetuses with open spinal dysraphism, there is a caudal displacement of the hindbrain and compression of the fourth ventricle and the CM. This is evidenced by a loss of the normal visualization of the IT or an IT-value under 1.5 mm (5). Moreover, due to the compression of fourth ventricle and CM, the brainstem-occipital bone distance (BSOB) appears smaller than the BS diameter itself, which is quantified by a BS/BSOB ratio ≥1 (Figure 2) (6).

In cases of cystic posterior fossa (such as DWM or Blake's pouch cyst), there is visualization of only two spaces rather than three. Moreover, the second space or fourth ventricle-cisterna magna complex, represents in this case the BSOB, which has been found to be enlarged in comparison to the BS diameter (7; 8). This results in a BS/BSOB ratio <1 (*Figure 3*) (7).

In summary, visualization of the three hypoechoic areas of the posterior brain at midsagittal view, together with its quantification (i.e., IT and BS/BSOB measurement) would allow the obstetrician to accurately rule out possible posterior brain malformations.

Rationale

Visualization of the posterior fossa brain spaces, their spatial relationship and measurements can be obtained in the midsagittal view of fetal head, the same used for NT measurement (9), and plays an important role in the early diagnosis of neural tube defects, such as open spinal dysraphism (5), and posterior fossa anomalies, such as DWM or BPC (7). However, assessment of the fetal posterior fossa in the first trimester is still challenging due to several limitations including involuntary movements of the fetus and small size of the brain structures, causing difficulties for examination and misdiagnosis. Moreover, it is also operator-dependent for the acquirement of high-quality ultrasound images, standard measurements, and precise diagnosis.

The use of new technologies to improve the acquisition of images, to help automatically perform measurements, or aid in the diagnosis of fetal abnormalities, may be of great importance for the optimal assessment of the fetal brain, particularly in the first trimester (10). Artificial intelligence (AI) is described as the ability of a computer program to perform processes associated with human intelligence, such as learning, thinking and problem-solving. Deep Learning (DL), a subset of Machine Learning (ML), is a branch of AI, defined by the ability to learn features automatically from data without human intervention. In DL, the input and output are connected by multiple layers loosely modeled on the neural pathways of the human brain. In the image recognition field, one of the most promising type of DL networks is represented by convolutional neural networks (CNN). These are designed to extract highly representative image features in a fully automated way, which makes them applicable to diagnostic decision-making.

Promising medical AI applications are emerging in the areas of screening, prediction, triage, diagnosis, drug development, treatment, monitoring and imaging interpretation (11). In the last decade, automatic measurements, and assessments based on artificial intelligence (AI) have been introduced in obstetric ultrasound, as an effort to reduce intra- and inter-observer measurement variations and improve diagnostic accuracy (12). The application of AI in obstetric ultrasound includes three aspects: structure identification, automatic and standardized measurements, and classification diagnosis (13). Since obstetric ultrasound is time-consuming, the use of AI could also reduce examination time and improve workflow (14).

Aim of the Study

According to these observations, we propose a research project aimed to develop an ultrasound-based AI-algorithm, which is capable to assess the fetal posterior fossa structures during the first trimester ultrasound scan and discriminate between normal and abnormal findings through a fully automatic data processing.

Methods

Study design: this is a multicenter retrospective observational cohort study and subsequent prospective cohort study. The study design will be organized in two different phases.

The **first phase**, the feasibility retrospective study, has the objective to develop and train AI-Algorithm with normal and abnormal images retrospectively acquired during first trimester ultrasound scan from ten international fetal medicine centers.

The **second phase,** a prospective clinical validation, has the objective to test the AI-Algorithm in the assessment of the fetal posterior fossa anatomy in a real clinic setting with real patients from each of the participating fetal medicine centers.

Setting: Three (3) fetal medicine centers.

Participants: singleton pregnant population who underwent ultrasound examination between 11 - 14 weeks of gestation in ten fetal medicine centers.

Inclusion Criteria: Women with single pregnancies who underwent ultrasound examination between 11+0 – 13+6 weeks of gestation or a fetal crown-rump-length between 45 – 84 mm.

Exclusion Criteria:

1. Women who did not have the first trimester screening scan at the settled gestational age.
2. Women in which a good visualization of the mid-sagittal view of the fetal head was not technically possible.
3. Women who are not able to give the informed consent.

ENDPOINTS

Primary endpoint: To validate a novel AI-based technology, which could potentially be used as a screening tool for fetal brain abnormal findings in the first trimester.

Secondary endpoints: To improve the performance of the standard first trimester screening of fetal posterior fossa ensuring its reliable sonographic assessment within a shorter time of execution. To detect higher repeatability and reproducibility, allowing to implement the ultrasound screening also in terms of efficiency on a vast scale, optimizing healthcare resources

Variables

Ultrasound images of the mid-sagittal view of the fetal head between 11+0 – 13+6 weeks of gestation or a fetal crown-rump-length between 45 – 84 mm, as described by the Fetal Medicine Foundation (FMF) for the NT measurement (Fig. 1). The magnification of the image should be such that the fetal head and one third of the thorax occupy the whole screen. A mid-sagittal view of the fetal head is defined by the presence of the echogenic tip of the nose and rectangular shape of the palate anteriorly, the translucent diencephalon in the center and the nuchal membrane posteriorly. The fetus should be in a neutral position, with the head in line with the spine (Fig 1).

Firstly, the images will be reviewed from data scientists, specialized in medical image analysis to assess image quality and verify that each image complies with the prescribed parameter settings (e.g., depth, focus, etc.), while those that do not meet the criteria will be discarded. Subsequently, images will undergo a pre-processing consisting of normalization and resizing. Resulting images will be randomly partitioned into "training-set" (85% of total cases) and "test-set" (15% of total cases). Data will be partitioned in order to have a balanced distribution of normal and abnormal cases, both in the training- and test-set.

Data Collection:

In the **first phase** of the study, participating fetal medicine centers will search their electronic databases for midsagittal images of singleton pregnant women who underwent ultrasound imaging at 11+0 – 13+6 weeks of gestation with any fetal posterior fossa anomaly, such as open spinal dysraphism, DWM or BPC. Normal images of the fetal posterior fossa at the same gestational age will be provided by the promoting centers – i.e., Fondazione Policlinico A. Gemelli, IRCCS and University of Parma. Clinical, ultrasound, prenatal and postnatal information of each case will be retrieved from patient's medical records and entered an electronic database collection file by the principal investigator from each participating center. The acquired images will be anonymized, saved as DICOM and shared through a dedicated cloud storage system which will be set up by the bioengineering team. Each center will be able to access the web system using a personal ID and password.

In the **second phase** of the study, the algorithm will be prospectively tested and validated in a real clinical setting with real patients from each of the participating fetal medicine centers. Inclusion and exclusion criteria, imaging protocol and data collection will be the same carried out during the retrospective phase.

Sample size: As this is a large-scale study, no formal sample size calculation is required. For the retrospective part of the study the centers will provide images of normal fetuses in a ratio of 1 to 1 in respect of the abnormal cases so we aim to collect at least 100 pathological images and 100 normal images in order to have a representative sample able to test the algorithm. For the prospective part of the study the number of patients to be recruited is expected to be around 10.000 in a ratio of 2 to 8 of the abnormal cases.

Study timeline: 3 years (36 months).

Year 1: Months 1 to 3: study set-up.

Month 3 to 6: retrospective data collection and image selection.

Month 6 to 12: image labeling and classification, retrospective data analysis, and AI-algorithm set-up and development.

Year 2: Month 13 to 24; ***Year 3***: Month 25 to Month 30: prospective data collection.

Year 2: Month 18 to 24; ***Year 3***: Month 25 to 32: analysis for artificial intelligence-algorithm validation, assessing the performance of the AI-based algorithm and investigating the AI-based algorithm repeatability and reproducibility.

Year 3 Month 32 to 36: Data analysis; conduction of main statistical analysis and compilation of final report; interpretation of results and manuscript (s) writing; final report, concluding remarks and plans.

Artificial intelligence methodology and statistical analysis:

For the aim of the study, two different AI methodologies will be used. Firstly, we aim to the definition and the tuning of the image processing algorithm to quantify patterns on the region of interest's grey level texture, which are descriptive of open spinal dysraphism, DWM or BPC. The algorithm will be fine-tuned to identify in such region, pre-defined patterns expressed in terms of high intensity versus low intensity areas, and to cluster homogenous grey level zones bounded by steep grey level gradients. To this purpose, the images will be processed through convolutional and wavelet filters, and then pixels of the region of interest will be assigned to one of the different clusters. The results of the algorithm in terms of pixel assignments will be evaluated with standard metrics. Spatial features of these pixel clusters, such as the elongation along the main orthogonal axis, will be computed and used as global image features.

Secondly, the dedicated CNN will be designed, in order to capture the composition of complex features, such as shapes and edges, decode image data and provide an accurate classification of the fetal posterior fossa appearance. The network architecture will consist of an input layer, receiving pre-processed (normalized and resized) image data, 10 to 20 hidden layers, extracting patterns within data, and an output layer, providing the classification.

Transformation parameters, such as rotation angle or the Gaussian filter standard deviation, will be randomly chosen. Regarding the cross-validation, we will use the k-fold cross-validation: the training set will be partitioned into k sub-groups, or folds, and the algorithm will be trained on k-1 folds and validated on the remaining fold, iterating the process k times; the parameter k will be set equal to 5. This strategy will overcome the lack of a huge amount of data, at the same allowing the algorithm to learn from the whole training set. The performance of the trained algorithm will be tested over the test set by using "ad hoc" metrics: the precision-recall (PR) curve and the Area Under the Curve (AUC) of the ROC curve for estimating the classification accuracy. Moreover, the Cohen's kappa will be calculated to evaluate the agreement between the output provided by the software and the gold standard, i.e. the classification provided by the experienced clinicians.

All analyses will be performed by using R software version 4.1.2 (CRAN ®, R Core Team, 2020).

REGULATORY ISSUES

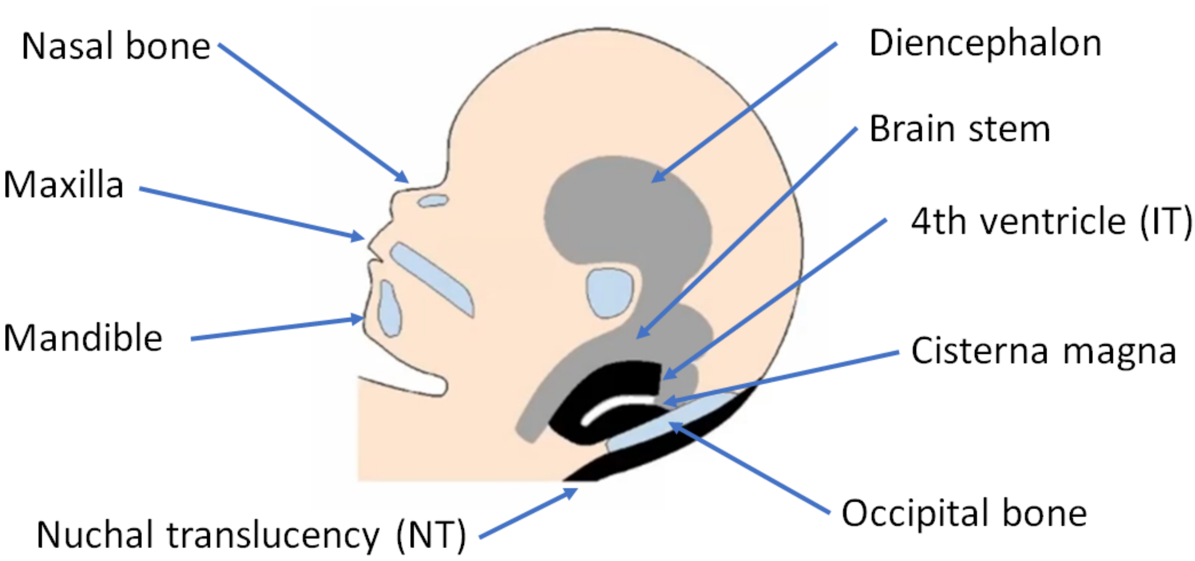
Ethics approval: The Study Coordination Centres will have to obtain approval from the local Ethics Committee (EC). The study will be conducted in accordance with the recommendations for physicians involved in research on human subjects adopted by the 18th World Medical Assembly, Helsinki 1964 and later revisions.

Consent: consent to enter the study will be sought from each participant only after a full explanation has been given, an information leaflet offered, and time allowed for consideration. Signed participant consent will be obtained. The right of the participant to refuse to participate without giving reasons will be respected. All participants are free to withdraw at any time from the protocol treatment without giving reasons and without prejudicing further treatment. In the other centres, each PI will use the appropriate way to consent women according to their local research ethics rules.

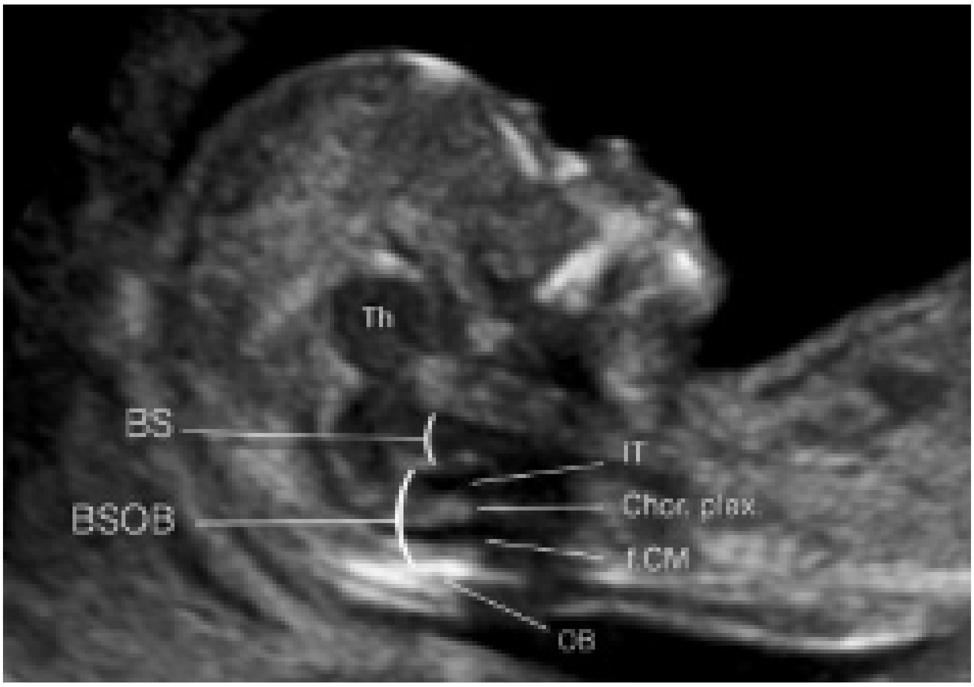
Figures

Figure 1. a) Schematic representation of the midsagittal view of the fetal brain on the first trimester ultrasound scan. B) sonographic representation of mid-sagittal view of fetal head for evaluating posterior fossa brain region.

a)



b)



Th: Thalamus; BS: Brainstem; IT: fourth ventricle or intracranial translucency;

Chor. Plex: posterior border of the IT is the choroid plexus of the fourth ventricle

BSOB: brainstem-occipital bone distance

Figure 2. Mid-sagittal view of fetal head in a case of open spina bifida: loss of the normal visualization of IT and increased ratio (≥1) between brainstem distance to brainstem-occipital bone (BSOB) distance

Figure 3. Mid-sagittal view of fetal head in a case of open spina bifida (left side) and Dandy Walker malformation (right side).

Rome, 15th January 2023

Dott.ssa Alessandra Familiari

References

1. ISUOG Practice Guidelines: performance of first-trimester fetal ultrasound scan. Guidelines, ISUOG. s.l. : Ultrasound Obstet Gynecol , 2013, Vols. 41: 102–113.

2. European network of population-based registries for the epidemiological surveillance of congenital anomalies. [Online] https://eu-rd-platform.jrc.ec.europa.eu/eurocat_en.

3. Outcome of fetal cerebral posterior fossa anomalies. . Long A, Moran P, Robson S. s.l. : Prenat Diagn. , 2006, Vols. Aug;26(8):707-10.

4. First-trimester cystic posterior fossa: reference ranges, associated findings, and pregnancy outcomes. . Garcia-Rodriguez R, Garcia-Delgado R, Romero-Requejo A, Medina-Castellano M, Garcia-Hernandez JA, Gonzalez-Martin JM, Sepulveda W. s.l. : J Matern Fetal Neonatal Med. , 2021, Vols. Mar;34(6):933-942.

5. Assessment of intracranial translucency (IT) in the detection of spina bifida at the 11-13-week scan. . Chaoui R, Benoit B, Mitkowska-Wozniak H, Heling KS, Nicolaides KH. s.l. : Ultrasound Obstet Gynecol., 2009, Vols. Sep;34(3):249-52.

6. Lachmann, Robert, et al. Posterior brain in fetuses with open spina bifida. Prenatal Diagnosis. 2011, pp. 31: 103-106.

7. Prospective detection and differential diagnosis of cystic posterior fossa anomalies by assessing posterior brain at 11e14 weeks. Volpe P, Persico N, Fanelli T, De Robertis V, D'Alessandro J, Boito S, Pilu G, Votino C. s.l. : Am J Obstet Gynecol MFM., 2019, Vols. May;1(2):173-181. .

8. Fetal Cystic Malformations of the Posterior Fossa in the First Trimester of Pregnancy. Nizard J, Bernard JP, Ville Y. s.l. : Fetal Diagn Ther., 2005, Vols. Mar-Apr;20(2):146-51. .

9. Appearance of the fetal posterior fossa at 11 + 3 to 13 + 6 gestational weeks on transabdominal ultrasound examination. Egle D, Strobl I, Weiskopf-Schwendinger V,

Grubinger E, Kraxner F, Mutz-Dehbalaie IS, Strasak A, Scheier M. s.l. : Dec;38(6):620-4., 2011, Vol. Ultrasound Obstet Gynecol. .

10. Machine Learning in Fetal Cardiology: What to Expect. . Garcia-Canadilla P, Sanchez-Martinez S, Crispi F, Bijnens B. s.l. : Fetal Diagnosis Ther, 2020, Vols. 47(5):363-372. .

11. Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology. Drukker L, Noble JA, Papageorghiou AT. s.l. : Ultrasound Obstet Gynecol., 2020, Vols. Oct;56(4):498-505.

12. Artificial Intelligence in Obstetric Ultrasound: An Update and Future Applications. Chen Z, Liu Z, Du M, Wang Z. s.l. : Front Med, 2021, Vol. Aug 27;8:733468.

13. Analysis of maturation features in fetal brain ultrasound via artificial intelligence for the estimation of gestational age. . Burgos-Artizzu XP, Coronado-Gutiérrez D, Valenzuela-Alcaraz B, Vellvé K, Eixarch E, Crispi F, Bonet-Carne E, Bennasar M, Gratacos E. s.l. : Am J Obstet Gynecol MFM. , 2021, Vol. Aug 14;3(6):100462. .

14. Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology. Drukker L, Noble JA, Papageorghiou AT. s.l. : Ultrasound Obstet Gynecol. , 2020, Vols. Oct;56(4):498-505.