

Machine Learning in Fetal Cardiology: What to Expect

Patricia Garcia-Canadilla^{a, b} Sergio Sanchez-Martinez^a Fatima Crispi^{a, c}
Bart Bijmens^{a, d, e}

^aInstitut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain; ^bInstitute of Cardiovascular Science, University College London, London, UK; ^cFetal Medicine Research Center, BCNatal – Barcelona Center for Maternal-Fetal and Neonatal Medicine (Hospital Clínic and Hospital Sant Joan de Déu), Institut Clínic de Ginecologia Obstetricia i Neonatologia, Centre for Biomedical Research on Rare Diseases (CIBER-ER), Barcelona, Spain; ^dDepartment of Cardiovascular Sciences, KU Leuven, Leuven, Belgium; ^eICREA, Barcelona, Spain

Keywords

Machine learning · Deep learning · Artificial intelligence · Fetal cardiology · Obstetrics · Echocardiography · Decision support systems

Abstract

In fetal cardiology, imaging (especially echocardiography) has demonstrated to help in the diagnosis and monitoring of fetuses with a compromised cardiovascular system potentially associated with several fetal conditions. Different ultrasound approaches are currently used to evaluate fetal cardiac structure and function, including conventional 2-D imaging and M-mode and tissue Doppler imaging among others. However, assessment of the fetal heart is still challenging mainly due to involuntary movements of the fetus, the small size of the heart, and the lack of expertise in fetal echocardiography of some sonographers. Therefore, the use of new technologies to improve the primary acquired images, to help extract measurements, or to aid in the diagnosis of cardiac abnormalities is of great importance for optimal assessment of the fetal heart. Machine learning (ML) is a computer science discipline focused on teaching a computer to perform tasks with specific goals without explicitly programming the rules on how to perform this task. In this re-

view we provide a brief overview on the potential of ML techniques to improve the evaluation of fetal cardiac function by optimizing image acquisition and quantification/segmentation, as well as aid in improving the prenatal diagnoses of fetal cardiac remodeling and abnormalities.

© 2020 S. Karger AG, Basel

Introduction

Fetal echocardiography was introduced to assess fetal cardiac function only 15 years ago (the first study was performed in 2004). It has evolved from the description of cardiac anatomical abnormalities toward quantitative assessment of cardiac dimensions, shape, and function and has been demonstrated to be useful in the diagnosis and monitoring of fetuses with a compromised cardiovascular system related to several fetal conditions, such as intrauterine growth restriction (IUGR), twin-to-twin transfusion syndrome, and congenital heart disease [1–3]. Moreover, some cardiac parameters have already shown to be helpful in predicting perinatal problems and long-term cardiovascular outcomes [4].

Different ultrasound (US) approaches are currently used to evaluate fetal cardiac function, including conven-

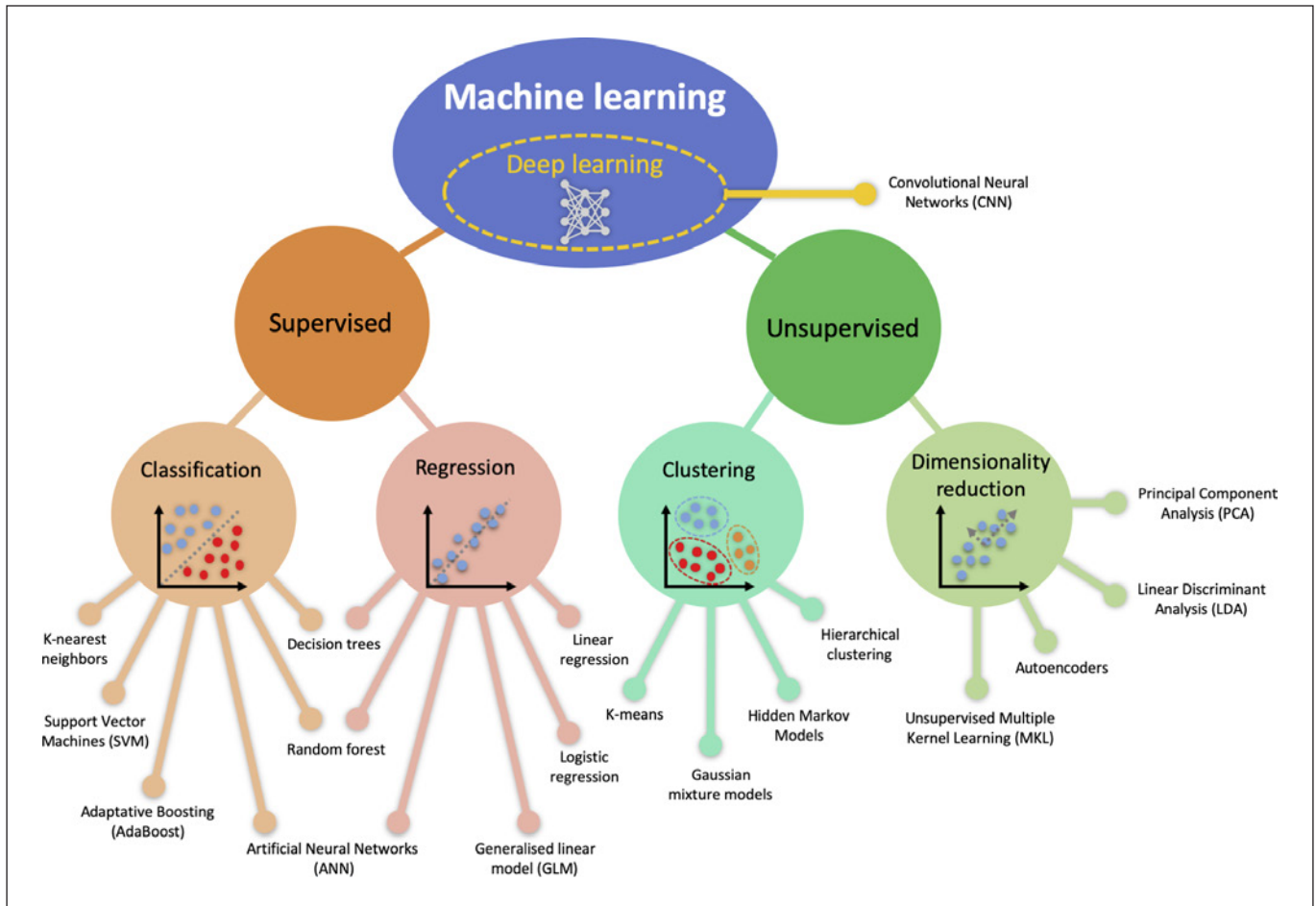


Fig. 1. Classification of ML/DL algorithms. DL is a subset of ML based on ANN and can be applied in a supervised or unsupervised manner.

tional 2-D imaging, M-mode, blood-pool and tissue Doppler imaging, 2-D speckle tracking, and 4-D spatiotemporal imaging correlation [4, 5]. For any evaluation, an optimal image of the fetal heart is crucial to adequately assess cardiac structure and function. However, assessing fetal cardiac function is still challenging due to involuntary movements of the fetus, the small size of the heart, the high heart rate, the limited access to the fetus, and the lack of expertise in fetal echocardiography of some sonographers. After having obtained an optimal image, measurements have to be performed in order to extract relevant cardiac features that relate to remodeling and functional status. Currently, these are mainly carried out manually by the sonographer, either during the investigation or offline using a dedicated workstation. Therefore, the use of new technologies to improve the primary acquired images or help extract and standardize measurements is of great importance for optimal assessment of the fetal heart.

Machine learning (ML) is a computer science discipline focused on teaching a computer to perform tasks, with a specific goal in mind, without explicitly programming the rules on how to perform this task. Mathematically speaking, learning occurs when a computer iteratively improves its performance on the given task (e.g., classification of a disease or estimation of clinical measurements) with experience or, in other words, when it is exposed to data [6]. Usually, ML algorithms are classified into 2 approaches: supervised and unsupervised learning algorithms (Fig. 1). Deep learning (DL), a popular algorithm (and often thought of when the term machine learning is used) is just a subset of ML that uses a layered structure of calculations known as artificial neural networks (ANN) on unstructured data. Figure 2 illustrates the typical pipeline for both supervised and unsupervised learning algorithms. Supervised learning requires explicit ground truth goals (diagnostic labels,

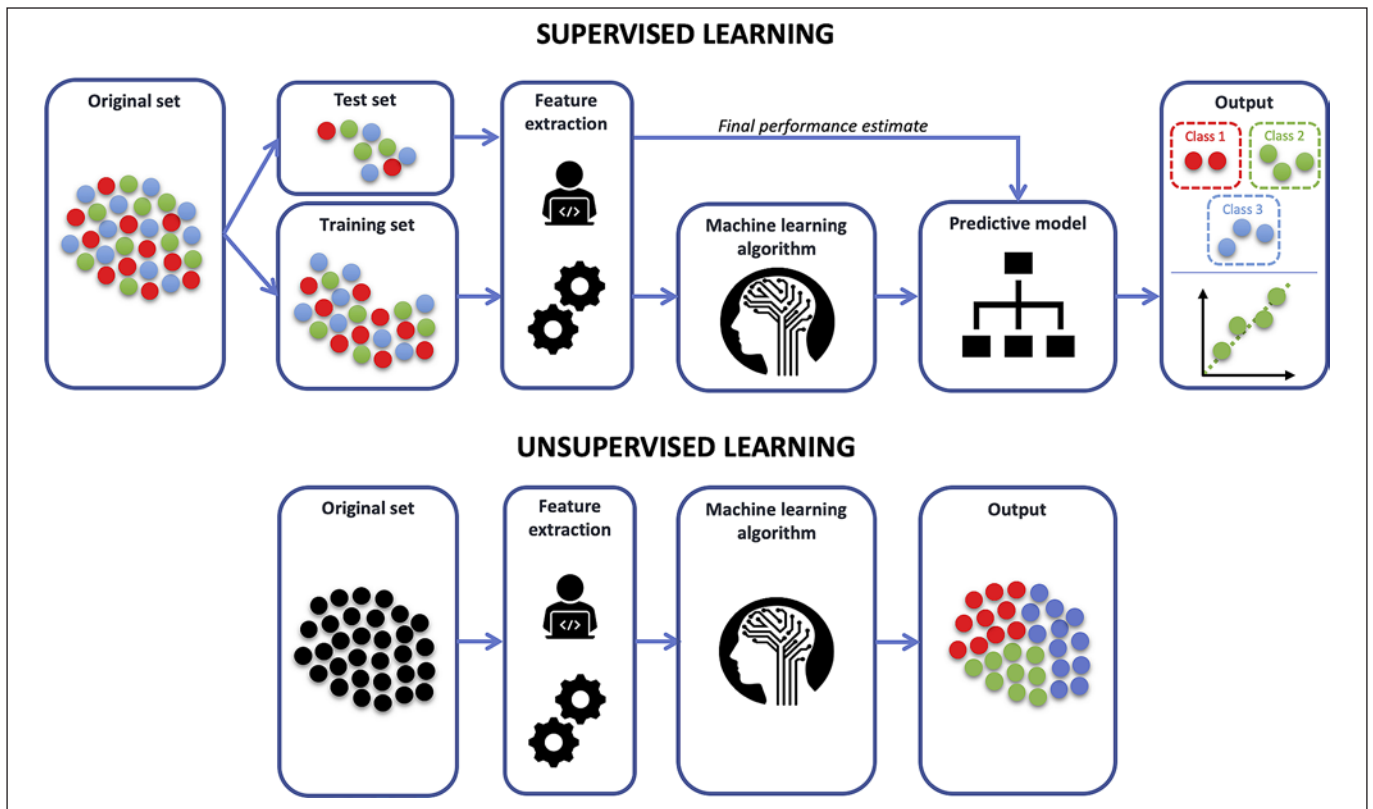


Fig. 2. Pipeline of supervised (top) and unsupervised (bottom) learning applications.

outcomes, reference image measurements, etc.) from which the algorithm can optimize its performance during training. Supervised learning algorithms can be further classified into classification and regression (Fig. 1). Classification techniques evaluate the given input and come up with a category such as “red” or “blue” or “disease” or “nondisease,” while regression techniques result in a continuous output: the value of the predicted quantity (such as the probability of a diagnosis). Besides DL, the most common classification algorithms include decision trees, support vector machines (SVM), etc., while linear and logistic regressions are typical regression algorithms (Fig. 1). On the other hand, unsupervised learning algorithms receive unlabeled examples and aim at discovering main patterns or similarities in the data, which would correspond to different disease manifestations or different phenotypes within a given disease, or different temporal evolution. Consequently, supervised learning is commonly used when the final goal is well known at the time of learning and unsupervised learning is used as an exploratory tool and usually the final goal follows from the analysis of the obtained

results. Unsupervised learning algorithms can be further classified into clustering and dimensionality reduction as illustrated in Figure 1. Typical clustering algorithms include K-means or Gaussian mixture models, while principal component analysis and linear discriminant analysis are classical dimensionality reduction techniques.

Once ML models are trained, their performance on unseen data (referred to as the test set) is known as the model’s generalizability (Fig. 2). Models that perform considerably better on the training set compared to the test set are overfitted, which means that they have a strong adherence to the training cases, but new patients are not correctly handled. Finding a good balance between training and testing performance is thus crucial for the application of ML models in clinical settings. A related highly relevant risk when using ML for clinical decision making is how to deal with, and not miss, rare occurrences in the (testing) data that were underrepresented in the training dataset. To circumvent this risk, ML approaches (especially supervised ones) need to be trained with a dataset that sufficiently captures the phenomenon under study.

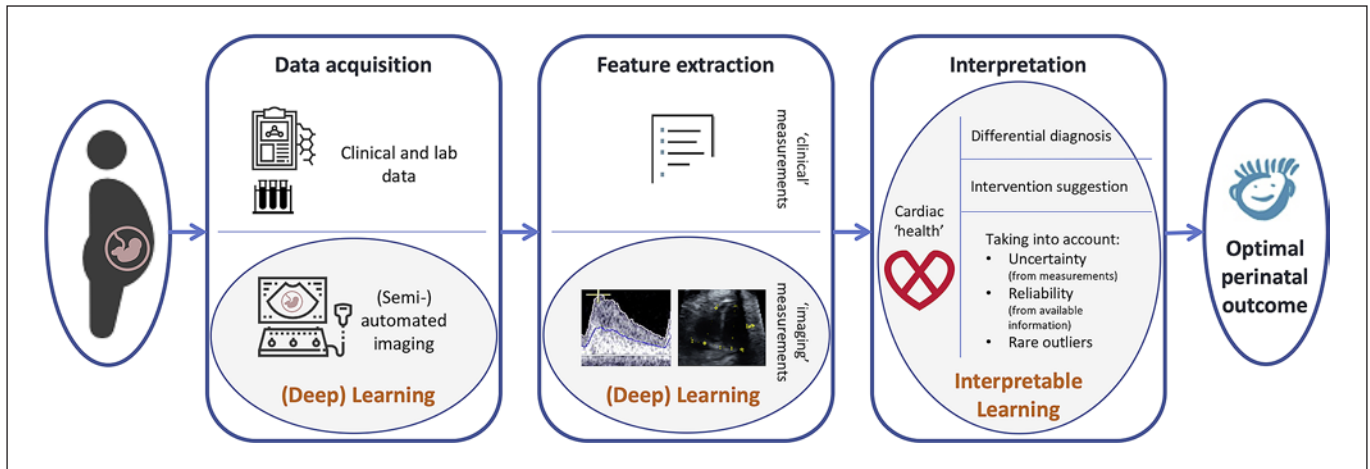


Fig. 3. Overview of the application of ML/DL in fetal cardiology. In the short term, ML will help the sonographer to acquire a full set of optimal and standardized images in the shortest possible time. This can improve data quality and interpretability. Next, ML will be used for (semi-) automated extraction of features (measurements) from the images, thus again improving the standardization and efficiency of the imaging department. This first component, essential for the use of images in fetal cardiology, will likely benefit greatly from the development in DL. The next step in clinical deci-

sion making is data interpretation for diagnosis and therapy planning. This component is much riskier so that interpretability and reliability of the ML decision support becomes a crucial factor. Therefore, in the foreseeable future, this will stay fully in the hands of the clinician but ML can provide helpful support by presenting the data in such a way that comparison of an individual patient with knowledge from pathophysiology and clinical trials/research becomes an easier task when a huge amount of complex data (from anamnesis to images over lab results) is available.

For clinical decision making, an unsupervised approach that highlights these rare instances might therefore be better compared to a supervised one that forces decisions towards what was trained for. In order to learn more about ML concepts, we refer the reader to the review paper by Deo [7].

ML techniques can help to optimize image acquisition protocols, thus reducing the acquisition time and ensuring optimal quality, and they can help to extract comprehensive and standardized information for a better evaluation of cardiac function. In this review we provide a brief overview on ML/DL applications in obstetrics, with a particular focus on the evaluation of fetal cardiac function by optimization of image acquisition and quantification/segmentation, and aid in improving the prenatal diagnoses of fetal cardiac remodeling and abnormalities.

ML for Data Acquisition

Image acquisition is the first step towards building a system to optimize the characterization of fetal cardiac function. This step is of capital importance, as the extracted information will be greatly conditioned by the intrinsic

quality and amount of input data. Acquisition of the best standard fetal views is labor intensive and relies on the sonographer's experience. The resulting interoperator variability in image acquisition hampers individual temporal follow-up or the combination of different data sources for research purposes. In this sense, ML-powered acquisition methods to speed up the acquisition, decrease the learning curve, and standardize the resulting images seem highly desirable, as they promise to boost data quality and standardization with minimal human intervention.

The improvement of image acquisition using ML is based on evaluation of the current (2-D/3-D) image on the screen by scoring how closely it resembles the type of view that was intended. This view was learned during a training phase (without explicitly defining the image appearance or content; this is learned by the algorithm). Many ML approaches can be used, but DL, using ANN, seems the most promising.

The acquisition of the fetal facial standard plane is a requisite for extracting biometric measurements and making a diagnosis during US examination. Lei et al. [8] automated this task with a SVM classifier. More recently, Yu et al. [9] leveraged the power of deep convolutional neural networks (CNN) to automatically recognize the

fetal facial standard plane during routine US examination. Another standard plane is that of the fetal abdominal region, which allows measurement of the abdominal circumference (AC) and estimation of the fetal weight as a proxy for fetal health. CNN have already been trained to automatically find the abdominal region in a US image and then determine the image quality by assessing the goodness of depiction for key structures such as the stomach bubble and the umbilical vein [10]. In a similar fashion, Rahmutallah et al. [11, 12] trained an adaptive boosting (AdaBoost) model to detect these 2 structures in 2-D US images for the purpose of scoring image quality. Other ensemble approaches have been proposed to categorize unlabeled fetal 2D US images. In particular, Yaqub et al. [13] used a random forest (RF) classifier to detect meaningful structures from different regions inside the images. A more ambitious project using CNN targeted the classification of a broader collection of fetal images planes by automatic recognition of 14 different fetal structures in 2-D US images [14]. Concerning 3-D fetal US, Raynaud et al. [15] proposed an ensemble of DL for feature extraction and RF for classification of organs with the purpose of automatically encoding anatomical variability while discarding the fetus pose.

Detection of the standard scan plane in fetal brain US is an essential step in the assessment of fetal development. This task was achieved by Li et al. [16] using a CNN approach in 3-D fetal US. Concerning quality control, Yaqub et al. [17] proposed a DL solution that automatically assessed whether transventricular 2-D US images of the fetal brain met clinical standards. Namely, they first localized the fetal brain, then detected the regions of interest, and finally learned the US patterns that enable plane verification. ML techniques have also been used to automatically identify the transthalamic plane in 3-D US to then assess brain biometrics such as the fetal biparietal diameter and head circumference (HC) [18].

Specific studies involving ML techniques for imaging the fetal heart are still scarce. Among the few examples found, Bridge et al. [19] implemented a framework for tracking the key variables appearing in freehand 2-D US scanning videos of the healthy fetal heart through the use of regression forests. Concerning the electrical activity, Yu et al. [20] used independent component analysis and Muduli et al. [21] used DL to reconstruct the fetal electrocardiogram from abdominal ECG recordings. A next step towards automating the fetal US scanning consists of coupling the image plane/volume recognition with a robot arm that performs the scanning, which was pioneered by Wang et al. [22].

ML approaches for improved fetal data acquisition are already a reality in research settings and they are expected to become clinically available in the short term (5 years). In the midterm, ML techniques may be combined with robotics to automatically extract standardized fetal imaging views.

ML for Image Quantification and Feature Extraction

Fetal biometric parameters such as HC, biparietal diameter, AC, femur length, and thickness of nuchal translucency are commonly used for estimation of fetal weight, gestational age (GA) and detection of fetal abnormalities during prenatal US examinations. An accurate estimation of fetal weight and GA is essential to detect any abnormal fetal growth pattern, such as small or large for GA, IUGR, or cardiac abnormalities. Kim et al. [23] recently published a DL model to automatically calculate the HC together with the biparietal diameter from 2-D US images. A different approach was used by Li et al. [24], who first used RF to localize the fetal head and then ellipse fitting to estimate the HC from 2-D US images. van Den Heuvel et al. [25] went a step further and implemented a DL model that calculated HC from obstetric sweep protocol data. These data likely do not contain standard planes, and thus their method has great potential for application in resource-constrained countries, where there is a lack of skilled obstetricians. Lorenz et al. [26] recently published a pipeline combining RF, shape models, and CNN to automatically perform view recognition and anatomical landmark location, with the objective of measuring the AC from 3-D US recordings. Similarly, Kim et al. [27] used a CNN to estimate AC from 2-D US data. For further information on biometric measurements, we refer the reader to a recent review of automated techniques for the interpretation of fetal abnormalities [28].

ML methods have been proposed in the last decades to improve the estimation of GA in women with uncertain or unknown menstrual dates [29] and to improve the estimation of fetal weight during gestation. For example, Ashley et al. [30] explored whether data available at birth can be used to accurately predict the estimated fetal weight over the course of gestation using different ML methods such as RF or regression trees in a database of more than 10,000 normal and high-risk pregnancies. The authors found that ML algorithms estimate fetal weight better than other commonly used methods. Chuang et al. [31] developed an ANN to estimate fetal weight using

morphometric data from 991 fetuses, reporting a mean absolute error of 6.15%.

Apart from measuring fetal biometrics and estimating fetal weight, recent ML approaches have been geared towards segmentation to identify fetal structures and organs to timely find fetal abnormalities so that necessary action can be taken. Namburete et al. [32] used an RF classifier to segment cranial pixels in 2-D US images. More recently, Li et al. [33] used a DL approach to automatically segment the fetal body and amniotic fluid from 2-D US data. Other examples of DL for segmentation have targeted the fetal brain and lungs [34, 35] and these 2 organs plus the placenta and the maternal kidneys from magnetic resonance imaging [36]. Lastly, an ensemble of decision trees has been used to automatically segment fetal brain structures in 3-D US images [37].

Concerning the fetal heart, the bulk of research focuses on automatically measuring the heartbeat. Some examples are the detection of cardiac activity from a pre-defined free-hand US sweep of the maternal abdomen using a classification model [38], extraction of the fetal heart rate from cardiotocograms (CTG) using dimensionality reduction [39], or measuring fetal QRS complexes from maternal ECG recordings using ANN [40]. More recently, Sulas et al. [41] used ANN to detect heart beats from pulse-wave Doppler envelope signals extracted from B-mode videos. For more on measuring cardiac activity from fetal US using ML techniques, the reader might be interested in the review paper by Alnuaimi et al. [42].

The application of ML algorithms to extract features from fetal echocardiographic data is already being used in some high-end scanners, in particular for the calculation of pulsatility indices from peripheral blood flow recordings. This is expected to be translated to cardiac flows soon. In the midterm, these scanners will also estimate GA and assess fetal growth based on automatic extraction of the different biometric measurements discussed above.

ML for Fetal Diagnosis

Prenatal diagnosis of fetal abnormalities has greatly benefited from advances in US technology and, in the last years, also from the advances in ML. ML algorithms have been used in different applications within fetal US medicine such as to predict preterm births [43, 44], the risk of euploidy, trisomy 21, and other chromosomal aneuploidies [45] or prediction of perinatal outcomes on asymptomatic short cervical length [46] among others. Regarding fetal cardiology, one of the subfields in which ML has

been extensively applied in the last decades is improvement of the diagnosis of fetal hypoxia or acidemia based on the analysis of CTG. CTG is routinely used to record and monitor the fetal heart rate and uterine contractions during the antepartum and intrapartum periods in order to detect the symptoms of fetal distress as early as possible. In clinical practice, CTG traces are visually examined by clinicians and their interpretation is largely dependent on the clinician's expertise, leading to high inter- and intraobserver variability. Therefore, despite the existence of standardized guidelines, the accuracy and robustness of CTG to improve prenatal outcomes remain controversial. The use of ML to improve the predictive capacity of CTG recordings was first presented by Bassil et al. [47] in the late 1980s. Since then, several attempts have been made to increase the effectiveness of the automatic evaluation of CTG traced using different ML and DL methods including ANN, SVM, and RF among others. Most of the publications have used 2 different open-access CTG databases to evaluate their proposed ML algorithms, i.e., one from the University Hospital in Brno (Czech Republic), including 552 CTG recordings [48], and another from the University of Porto (Portugal), which includes 2,126 CTG recordings [49]. We have summarized the publications on the use of ML in the analysis of CTG for the last 10 years in online supplementary Table S1 (for all online suppl. material see www.karger.com/doi/10.1159/000505021). For a review of older publications, we refer the reader to the review of Graham et al. [50]. The best results were obtained by Ir-aji et al. [51] using the Portuguese database showing an accuracy of 99.5%. There have also been some attempts to translate this into clinical practice via the development of software such as Infant, PeriCALM [52, 53], and Foetos [54] or the development of mobile/website applications [55, 56] to provide additional support in the interpretation of CTG signals and therefore to improve the assessment of fetal status. However, there is no evidence on whether these systems really improve the prediction of fetal distress or acidemia compared to visual CTG interpretation alone, and reports about their clinical performance were not found. In a recent systematic review, the degree of interobserver reliability between human and ML interpretations of CTG signals was determined [57], and it was concluded that the use of ML for interpretation of CTG during labor does not improve neonatal outcomes and has yet to prove its reliability relative to expert observers. The root of the problem may be that any supervised ML-based system needs to be trained with human annotations and, given that the benefit of

CTG themselves for labor monitoring has not been clearly demonstrated, it is not surprising that adding an automatic system to evaluate CTG signals with similar information does not offer advantages in reducing adverse perinatal outcomes.

IUGR, which affects about 10% of pregnancies, has been associated with cardiac remodeling in utero that can persist postnatally [58–60]. Early detection of IUGR can improve the perinatal outcomes of these fetuses and reduce the risk of cardiovascular mortality in adulthood. The first study proposing the use of ML for the detection of IUGR using biometric data was presented by Gurgen et al. [61] in 1997. In that study an ANN was implemented to approximate the growth curves of fetuses showing an accuracy of 95% in the detection of IUGR. Later, Magenes et al. [62] proposed an SVM to detect IUGR using CTG data and showed good classification results in a cohort of 70 fetuses. In 2014, Gadagkar et al. [63] developed an ANN system for the diagnosis of IUGR using only 2-D US morphometric measurements from almost 300 fetuses, and they had results similar to those obtained clinically in the same study population. Similarly, Rawat et al. [64] implemented an ANN model using again 2-D US morphometric measurements from a total of 120 fetuses. Recently, Kuhle et al. [65] compared different ML methods to predict fetal growth abnormalities in a cohort of more than 30,000 patients. However, the authors reported that the ML methods used did not offer any advantage over logistic regression in the prediction of fetal growth abnormalities. The main limitation of all of these studies is that the detection of IUGR was performed considering only morphometric data, which only provide information about the fetal weight, without considering any other data such as blood flow velocities or cardiac deformation measured by Doppler or B-mode US, respectively. It is known that IUGR fetuses show abnormal blood flow patterns in the fetal circulation detected by Doppler US [66, 67] and also signs of longitudinal systolic dysfunction [58]. It was recently demonstrated that unsupervised ML algorithms using both echocardiographic (including myocardial strain traces) and clinical data can be used to find groups of similar patients within a heart failure cohort and identify individuals with a beneficial response to cardiac resynchronization therapy [68]. A similar approach integrating clinical and heterogeneous echocardiographic data could be implemented to improve the detection of IUGR fetuses, identify those at high risk of adverse perinatal outcomes, and aid clinicians in finding optimal treatment strategies. However,

ML methods require a large number of patients during training in order to be able to capture the range of possible abnormalities, which is a limitation in fetal medicine as the number of patients is scarce. One possibility to overcome this limitation is to combine ML with “data augmentation” through physiological computational modelling as proposed by Hoodbhoy et al. [69]. Lumped models of the fetal circulation have demonstrated to be able to realistically simulate the hemodynamics of the fetus in many different conditions [66, 67, 70], thus providing virtual, but physiologically plausible, Doppler traces. Using these models, virtual patient populations can be created where the ratio of abnormal/normal cases can be increased so that the learning of the ML algorithms is less dependent on the data provided.

Finally, ML has been recently applied to improve the prenatal diagnosis of congenital heart diseases. Yeo et al. [71] presented an intelligent navigation method called FINE to automatically obtain different echocardiography anatomical views of the fetal heart and identify abnormalities within the cardiac anatomy. The tool was able to demonstrate evidence of abnormal fetal cardiac anatomy in 4 abnormal cases [71]. More recently, Arnaout et al. [72] proposed the use of a fully convolutional DL method in a supervised manner to: (1) identify the 5 most important views of the fetal heart, (2) segment and measure the cardiac structures, and (3) distinguish between normal hearts and tetralogy of Fallot and hypoplastic left heart syndrome using 685 echocardiograms from fetuses from 18 to 24 weeks of GA [72]. The best results were obtained in the diagnosis of hypoplastic left heart syndrome versus normality with a sensitivity and specificity of 100 and 90%, respectively. Although the results look promising, one of the main limitations of this study is that only 2 congenital heart diseases were evaluated and the DL system was only trained with images from 1 US machine without considering the variability in echocardiographs. Therefore, further studies with bigger datasets from different US machines need to be performed.

Conclusions

Given that ML approaches have become ubiquitous in our daily lives, they will become more and more integrated in clinical practice and in the assessment of the fetal heart. It is important to distinguish the different tasks involved in clinical decision making to understand how, and which type of, ML can be optimally employed. To obtain the best image quality in the shortest possible

time and with the smallest learning curve, as well as for standardized extraction of specific measurements from the images, ML approaches based on DL have shown great promise and are currently being implemented in high-end clinical scanners. However, when the diagnostic interpretation is performed, and especially when a treatment decision needs to be made, the “black-box” approach inherent to, for example, DL becomes problematic given its dependence on a large and very inclusive dataset with correct clinical labels and the inherent difficulty of providing an intuitive clinical explanation for the proposed decision. Here, other ML approaches, based on for example the identification of individuals with similar (complex and multimodal) clinical data and imaging features, seem more promising and are explored in different centers.

Therefore, when carefully used and validated, and taking into account all privacy, security, and auditing measures relevant for the use of clinical data, ML can play an important role in standardization of fetal cardiac data and provide support in the clinical interpretation and suggestion of the best preventive and interventional approach to optimize perinatal as well as long-term cardiovascular health (Fig. 3).

References

- Rychik J, Tian Z, Bebbington M, Xu F, McCann M, Mann S, et al. The twin-twin transfusion syndrome: spectrum of cardiovascular abnormality and development of a cardiovascular score to assess severity of disease. *Am J Obstet Gynecol*. 2007 Oct;197(4):392.e1–8.
- Crispi F, Hernandez-Andrade E, Pelsers MM, Plasencia W, Benavides-Serralde JA, Eixarch E, et al. Cardiac dysfunction and cell damage across clinical stages of severity in growth-restricted fetuses. *Am J Obstet Gynecol*. 2008 Sep;199(3):254.e1–8.
- Van Mieghem T, Gucciardo L, Doné E, Van Schoubroeck D, Graatsma EM, Visser GH, et al. Left ventricular cardiac function in fetuses with congenital diaphragmatic hernia and the effect of fetal endoscopic tracheal occlusion. *Ultrasound Obstet Gynecol*. 2009 Oct;34(4):424–9.
- Crispi F, Gratacs E. Fetal cardiac function: technical considerations and potential research and clinical applications. *Fetal Diagn Ther*. 2012;32(1–2):47–64.
- Crispi F, Valenzuela-Alcaraz B, Cruz-Lemini M, Gratacós E. Ultrasound assessment of fetal cardiac function. *Australas J Ultrasound Med*. 2013 Nov;16(4):158–67.
- Mitchell TM. *Machine Learning*. New York: McGraw-Hill; 1997.
- Deo RC. Machine learning in medicine. *Circulation*. 2015 Nov;132(20):1920–30.
- Lei B, Zhuo L, Chen S, Li S, Ni D, Wang T. Automatic recognition of fetal standard plane in ultrasound image: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). Piscataway: IEEE; 2014. p. 85–8.
- Yu Z, Wu L, Ni D, Chen S, Li S, Wang T, et al. Fetal facial standard plane recognition via deep convolutional neural networks. *Zhongguo Shengwu Yixue Gongcheng Xuebao*. 2017;36(3):267–75.
- Wu L, Cheng JZ, Li S, Lei B, Wang T, Ni D. FUIQA: fetal ultrasound image quality assessment with deep convolutional networks. *IEEE Trans Cybern*. 2017;47(5):1336–49.
- Rahmatullah B, Sarris I, Papageorgiou A, Noble JA. Quality control of fetal ultrasound images: Detection of abdomen anatomical landmarks using AdaBoost. *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE; 2011; pp 6–9.
- Rahmatullah B, Papageorgiou AT, Noble JA. Image analysis using machine learning: anatomical landmark detection in fetal ultrasound images. *2012 IEEE 36th Annual Computer Software and Applications Conference*. Piscataway: IEEE; 2012. p. 354–5.
- Yaqub M, Kelly B, Papageorgiou AT, Noble JA. Guided Random Forests for Identification of Key Fetal Anatomy and Image Categorization in Ultrasound Scans. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham; 2015. p. 687–94.
- Sridar P, Kumar A, Quinton A, Nanan R, Kim J, Krishnakumar R. Decision fusion-based fetal ultrasound image plane classification using convolutional neural networks. *Ultrasound Med Biol*. 2019 May;45(5):1259–73.
- Raynaud C, Ciofolo-Veit C, Lefèvre T, Ardon R, Cavallaro A, Salim I, et al. Multi-organ Detection in 3D Fetal Ultrasound with Machine Learning. *International Workshop on Fetal and Infant Image Analysis*. Springer, Cham; 2017; pp 62–72.
- Li Y, Khanal B, Hou B, Alansary A, Cerrrolaza JJ, Sinclair M, et al. Standard plane detection in 3D fetal ultrasound using an iterative transformation network. Cham: Springer; 2018.
- Yaqub M, Kelly B, Papageorgiou AT, Noble JA. A deep learning solution for automatic fetal neurosonographic diagnostic plane verification using clinical standard constraints. *Ultrasound Med Biol*. 2017 Dec;43(12):2925–33.

Disclosure Statement

The authors have no conflict of interests to declare.

Funding Sources

This project was partially funded by the “la Caixa” Foundation under grant agreement LCF/PR/GN14/10270005, the Instituto de Salud Carlos III (PI17/00675) integrados en el Plan Nacional de I+D+I y cofinanciados por el ISCIII-Subdirección General de Evaluación y el Fondo Europeo de Desarrollo Regional (FEDER) “Una manera de hacer Europa”, the Centro de Investigación Biomédica en Red de Enfermedades Raras (ERPR04G719/2016), the Cerebra Foundation for the Brain-Injured Child (Carmarthen, Wales, UK), and AGAUR 2017 SGR grant No. 1531.

Author Contributions

Patricia Garcia-Canadilla, Sergio Sanchez-Martinez, and Bart Bijmens participated in the conception and design of this review, drafted this work, approved the final version, and agreed on the accuracy and integrity of this work.

Fatima Crispi participated in the design of this review, revised it critically for important intellectual content, approved the final version, and agreed on the accuracy and integrity of this work.

- 18 Ambrose Grandjean G, Hossu G, Bertholdt C, Noble P, Morel O, Grangé G. Artificial intelligence assistance for fetal head biometry: assessment of automated measurement software. *Diagn Interv Imaging*. 2018 Sep;99(11):709–16.
- 19 Bridge CP, Ioannou C, Noble JA. Automated annotation and quantitative description of ultrasound videos of the fetal heart. *Med Image Anal*. 2017 Feb;36:147–61.
- 20 Yu Q, Yan H, Song L, Guo W, Liu H, Si J, et al. Automatic identifying of maternal ECG source when applying ICA in fetal ECG extraction. *Biocybern Biomed Eng*. 2018 Jan;38(3):448–55.
- 21 Muduli PR, Gunukula RR, Mukherjee A. A deep learning approach to fetal-ECG signal reconstruction. *2016 Twenty Second National Conference on Communication (NCC)*. IEEE; 2016; pp 1–6.
- 22 Wang S, Housden J, Noh Y, Singh D, Singh A, Skelton E, et al. *Robotic-assisted ultrasound for fetal imaging: evolution from single-arm to dual-arm system* [Internet]. 2019 Feb [cited 2019 Aug 21]. Available from: <http://arxiv.org/abs/1902.05458>
- 23 Kim HP, Lee SM, Kwon JY, Park Y, Kim KC, Seo JK. Automatic evaluation of fetal head biometry from ultrasound images using machine learning. *Physiol Meas*. 2019 May;40(6):065009.
- 24 Li J, Wang Y, Lei B, Cheng JZ, Qin J, Wang T, et al. Automatic fetal head circumference measurement in ultrasound using random forest and fast ellipse fitting. *IEEE J Biomed Health Inform*. 2018 Jan;22(1):215–23.
- 25 van den Heuvel TL, Petros H, Santini S, de Korte CL, van Ginneken B. Automated fetal head detection and circumference estimation from free-hand ultrasound sweeps using deep learning in resource-limited countries. *Ultrasound Med Biol*. 2019 Mar;45(3):773–85.
- 26 Lorenz C, Brosch T, Ciofolo-Veit C, Tobias K, Lefevre T, Salim I, et al. Automated abdominal plane and circumference estimation in 3D US for fetal screening. In: Angelini ED, Landman BA, editors. *Robotic-assisted ultrasound for fetal imaging: evolution from single-arm to dual-arm system*. SPIE; 2018; p. 17.
- 27 Kim B, Kim KC, Park Y, Kwon JY, Jang J, Seo JK. Machine-learning-based automatic identification of fetal abdominal circumference from ultrasound images. *Physiol Meas*. 2018 Oct;39(10):105007.
- 28 Rawat V, Jain A, Shrimali V. Automated techniques for the interpretation of fetal abnormalities: a review. *Appl Bionics Biomech*. 2018 Jun;2018:1–11.
- 29 Papageorgiou AT, Kemp B, Stones W, Ohuma EO, Kennedy SH, Purwar M, et al. Ultrasound-based gestational-age estimation in late pregnancy. *Ultrasound Obstet Gynecol*. 2016 Dec;48(6):719–26.
- 30 Naimi AI, Platt RW, Larkin JC. Machine learning for fetal growth prediction. *Epidemiology*. 2018;29(2):290–8.
- 31 Chuang L, Hwang JY, Chang CH, Yu CH, Chang FM. Ultrasound estimation of fetal weight with the use of computerized artificial neural network model. *Ultrasound Med Biol*. 2002 Aug;28(8):991–6.
- 32 Namburete AI, Noble JA. *Fetal cranial segmentation in 2D ultrasound images using shape properties of pixel clusters: 2013 IEEE 10th International Symposium on Biomedical Imaging*. Piscataway: IEEE; 2013. p. 720–3.
- 33 Li Y, Xu R, Ohya J, Iwata H. Automatic fetal body and amniotic fluid segmentation from fetal ultrasound images by encoder-decoder network with inner layers. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*. 2017; 1485–8.
- 34 Rajchl M, Lee MC, Oktay O, Kamnitsas K, Passerat-Palmbach J, Bai W, et al. DeepCut: object segmentation from bounding box annotations using convolutional neural networks. *IEEE Trans Med Imaging*. 2017;36(2):674–83.
- 35 Burgos-Artizzu XP, Perez-Moreno Á, Coronado-Gutierrez D, Gratacos E, Palacio M. Evaluation of an improved tool for non-invasive prediction of neonatal respiratory morbidity based on fully automated fetal lung ultrasound analysis. *Sci Rep*. 2019 Feb;9(1):1950.
- 36 Wang G, Li W, Zuluaga MA, Pratt R, Patel PA, Aertsen M, et al. Interactive medical image segmentation using deep learning with Image-specific fine tuning. *IEEE Trans Med Imaging*. 2018 Jul;37(7):1562–73.
- 37 Huang R, Xie W, Alison Noble J. VP-Nets: efficient automatic localization of key brain structures in 3D fetal neurosonography. *Med Image Anal*. 2018 Jul;47:127–39.
- 38 Maraci MA, Bridge CP, Napolitano R, Papa-georghiou A, Noble JA. A framework for analysis of linear ultrasound videos to detect fetal presentation and heartbeat. *Med Image Anal*. 2017 Apr;37:22–36.
- 39 Krupa N, Ma M, Zahedi E, Ahmed S, Hassan FM. MA M, Zahedi E, Ahmed S, Hassan FM. Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine. *Biomed Eng Online*. 2011 Jan;10(1):6.
- 40 M Lukoševičius, Marozas V. Noninvasive fetal QRS detection using Echo State Network. *Comput Cardiol* [Internet]. 2013 [cited 2019 Jan 24]. Available from: <https://ieeexplore.ieee.org/abstract/document/6712447>
- 41 Sulas E, Ortu E, Raffo L, Urru M, Tumbarello R, Pani D. *Automatic recognition of complete atrioventricular activity in fetal pulsed-wave Doppler signals*. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Piscataway: IEEE; 2018. p. 917–20.
- 42 Alnuaimi SA, Jimaa S, Khandoker AH. Fetal cardiac Doppler signal processing techniques: challenges and future research directions. *Front Bioeng Biotechnol*. 2017 Dec;5:82.
- 43 Rittenhouse KJ, Vwalika B, Keil A, Winston J, Stoner M, Kapasa M, et al. Improving preterm newborn identification in low-resource settings with machine learning. *PLoS One* 2019 Feb; 27:e0198919.
- 44 Hamilton EF, Dyachenko A, Ciampi A, Maurer K, Warrick PA, Garite TJ. Estimating risk of severe neonatal morbidity in preterm births under 32 weeks of gestation. *J Matern Fetal Neonatal Med*. 2020 Jan;33(1):73–80.
- 45 Neocleous AC, Nicolaides KH, Schizas CN. First trimester noninvasive prenatal diagnosis: a computational intelligence approach. *IEEE J Biomed Health Inform*. 2016 Sep;20(5):1427–38.
- 46 Bahado-Singh RO, Sonek J, McKenna D, Cool D, Aydas B, Turkoglu O, et al. Artificial Intelligence and amniotic fluid multiomics analysis: the prediction of perinatal outcome in asymptomatic short cervix. *Ultrasound Obstet Gynecol*. 2019 Jul;54(1):110–8.
- 47 Bassil HE, Dripps JH. Real time processing and analysis of fetal phonocardiographic signals. *Clin Phys Physiol Meas*. 1989;10:67–74.
- 48 Chudáček V, Spilka J, Burša M, Janků P, Hruban L, Hupčtych M, et al. *Open access intrapartum CTG database*. 2014 [cited 2019 Jun 13]. Available from: <http://www.biomed-central.com/1471-2393/14/16DATABASE>.
- 49 Ayres-de-Campos D, Sousa P, Costa A, Bernardes J. Omniview-SisPorto® 3.5 – a central fetal monitoring station with online alerts based on computerized cardiotocogram+ST event analysis. *J Perinat Med*. 2008 Jan;36(3):260–4.
- 50 Graham EM, Petersen SM, Christo DK, Fox HE. Intrapartum electronic fetal heart rate monitoring and the prevention of perinatal brain injury. *Obstet Gynecol*. 2006 Sep;108:656–66.
- 51 Iraj MS. Prediction of fetal state from the cardiotocogram recordings using neural network models. *Artif Intell Med*. 2019 May;96:33–44.
- 52 Brocklehurst P. A study of an intelligent system to support decision making in the management of labour using the cardiotocograph – the INFANT study protocol. *BMC Pregnancy Childbirth*. 2016 Dec;16(1):10.
- 53 Hamilton EF, Warrick PA. New perspectives in electronic fetal surveillance. *J Perinat Med*. 2013 Jan;41(1):83–92. [cited 2019 Aug 21].
- 54 Alonso-Betanzos A, Moret-Bonillo V, Hernandez-Sande C. Foetos: an expert system for fetal assessment. *IEEE Trans Biomed Eng*. 1991;38(2):199–211.
- 55 Akbulut A, Ertugrul E, Topcu V. Fetal health status prediction based on maternal clinical history using machine learning techniques. *Comput Methods Programs Biomed*. 2018 Sep;163:87–100.
- 56 Kazantsev A, Ponomareva J, Kazantsev P, Digilov R, Huang P. *Development of e-health network for in-home pregnancy surveillance based on artificial intelligence: proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*. Piscataway: IEEE; 2012. p. 82–4.

- 57 Balayla J, Shrem G. Use of artificial intelligence (AI) in the interpretation of intrapartum fetal heart rate (FHR) tracings: a systematic review and meta-analysis. *Arch Gynecol Obstet*. 2019;300(May):1–8.
- 58 Crispi F, Bijmens B, Figueras F, Bartrons J, Eixarch E, Le Noble F, et al. Fetal growth restriction results in remodeled and less efficient hearts in children. *Circulation*. 2010 Jun;121(22):2427–36.
- 59 Cruz-Lemini M, Crispi F, Valenzuela-Alcaraz B, Figueras F, Sitges M, Bijmens B, et al. Fetal cardiovascular remodeling persists at 6 months in infants with intrauterine growth restriction. *Ultrasound Obstet Gynecol*. 2016 Sep;48(3):349–56.
- 60 Sarvari SI, Rodriguez-Lopez M, Nuñez-García M, Sitges M, Sepulveda-Martinez A, Camara O, et al. Persistence of cardiac remodeling in preadolescents with fetal growth restriction. *Circ Cardiovasc Imaging*. 2017 Jan;10(1):e005270.
- 61 Gurgun F, Onal E, Varol FG. IUGR detection by ultrasonographic examinations using neural networks. *IEEE Eng Med Biol Mag*. 1997;16(3):55–8.
- 62 Magenes G, Pedrinazzi L, Signorini MG. Identification of fetal sufferance antepartum through a multiparametric analysis and a support vector machine: the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Piscataway: IEEE; 2004. p. 462–5.
- 63 Gadagkar AV, Shreedhara KS. Features Based IUGR Diagnosis Using Variational Level Set Method and Classification Using Artificial Neural Networks. 2014 Fifth International Conference on Signal and Image Processing. IEEE; 2014; pp 303–9.
- 64 Rawat V, Jain A, Shrimali V, Rawat A. Automatic detection of fetal abnormality using head and abdominal circumference. *Lecture Notes in Computer Science*. Cham: Springer; 2016. p. 525–34.
- 65 Kuhle S, Maguire B, Zhang H, Hamilton D, Allen AC, Joseph KS, et al. Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: a retrospective cohort study. *BMC Pregnancy Childbirth*. 2018 Dec;18(1):333.
- 66 Garcia-Canadilla P, Rudenick PA, Crispi F, Cruz-Lemini M, Palau G, Camara O, et al. A computational model of the fetal circulation to quantify blood redistribution in intrauterine growth restriction. *PLoS Comput Biol*. 2014 Jun;10(6):e1003667.
- 67 Garcia-Canadilla P, Crispi F, Cruz-Lemini M, Triunfo S, Nadal A, Valenzuela-Alcaraz B, et al. Patient-specific estimates of vascular and placental properties in growth-restricted fetuses based on a model of the fetal circulation. *Placenta*. 2015 Jul;36(9):981–9.
- 68 Cikes M, Sanchez-Martinez S, Claggett B, Duchateau N, Piella G, Butakoff C, et al. Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. *Eur J Heart Fail*. 2019 Jan;21(1):74–85.
- 69 Hoodbhoy Z, Hasan B, Jehan F, Bijmens B, Chowdhury D. Machine learning from fetal flow waveforms to predict adverse perinatal outcomes: a study protocol. *Gates Open Res*. 2018 Feb;2:8.
- 70 Gímenez Mínguez P, Bijmens B, Bernardino G, Lluch È, Soveral I, Gómez O, et al. Assessment of haemodynamic remodeling in fetal aortic coarctation using a lumped model of the circulation. In: Pop M, Wright G, editors. *Functional imaging and modelling of the heart*. Lecture Notes in Computer Science. Cham: Springer; 2017.
- 71 Yeo L, Romero R. Fetal Intelligent Navigation Echocardiography (FINE): a novel method for rapid, simple, and automatic examination of the fetal heart. *Ultrasound Obstet Gynecol*. 2013 Sep;42(3):268–84.
- 72 Arnaout R, Curran L, Chinn E, Zhao Y, Moon-Grady A. Deep-learning models improve on community-level diagnosis for common congenital heart disease lesions. 2018. DOI: arXiv:1809.06993v1.