Fetal electrocardiography and artificial intelligence for prenatal detection of congenital heart disease

Received: 19 April 2023 | Revised: 9 June 2023 | Accepted: 9 June 2023

DOI: 10.1111/aogs.14623

AOGS Acta Obstetricia et Gynecologica Scandinavica

ORIGINAL RESEARCH ARTICLE

Fetal electrocardiography and artificial intelligence for prenatal detection of congenital heart disease

Ivar R. de Vries[1,2] | Judith O. E. H. van Laar[1,2] | Marieke B. van der Hout-van der Jagt[1,2,3] | Sally-Ann B. Clur[4,5] | Rik Vullings[2,6]

[1]Department of Obstetrics and Gynecology, Máxima Medical Center, Veldhoven, The Netherlands

[2]Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

[3]Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

[4]Department of Pediatric Cardiology, Emma Children's Hospital, Amsterdam University Medical Centers, Academic Medical Center, Amsterdam, The Netherlands

[5]European Reference Network for rare, low prevalence and complex diseases of the heart - ERN GUARD-Heart (ERN GUARDHEART), Amsterdam, The Netherlands

[6]Nemo Healthcare BV, Veldhoven, The Netherlands

Correspondence
Ivar de Vries, Máxima Medical Center Veldhoven, De Run 4600, 5504 DB Veldhoven, The Netherlands.
Email: ivar.de.vries@mmc.nl

Funding information
ZonMw, Grant/Award Number: 10070022010001

Abstract

Introduction: This study aims to investigate non-invasive electrocardiography as a method for the detection of congenital heart disease (CHD) with the help of artificial intelligence.

Material and methods: An artificial neural network was trained for the identification of CHD using non-invasively obtained fetal electrocardiograms. With the help of a Bayesian updating rule, multiple electrocardiographs were used to increase the algorithm's performance.

Results: Using 122 measurements containing 65 healthy and 57 CHD cases, the accuracy, sensitivity, and specificity were found to be 71%, 63%, and 77%, respectively. The sensitivity was however 75% and 69% for CHD cases requiring an intervention in the neonatal period and first year of life, respectively. Furthermore, a positive effect of measurement length on the detection performance was observed, reaching optimal performance when using 14 electrocardiography segments (37.5 min) or more. A small negative trend between gestational age and accuracy was found.

Conclusions: The proposed method combining recent advances in obtaining non-invasive fetal electrocardiography with artificial intelligence for the automatic detection of CHD achieved a detection rate of 63% for all CHD and 75% for critical CHD. This feasibility study shows that detection rates of CHD might improve by using electrocardiography-based screening complementary to the standard ultrasound-based screening. More research is required to improve performance and determine the benefits to clinical practice.

KEYWORDS
artificial intelligence, congenital heart disease, fetal electrocardiography, fetal heart, prenatal diagnosis

Abbreviations: AUC, area under the ROC curve; CHD, congenital heart disease; ECG, electrocardiography; FHR, fetal heart rate; ROC, receiver-operating characteristics.

1 | INTRODUCTION

Congenital heart disease (CHD) is the most common severe congenital anomaly worldwide. With an incidence of eight per 1000 births,[1] around 1.35 million newborns are born with CHD—a third of which are severe[1–3] (i.e. requiring an intervention in the first year of life)—every year.[4] CHD is associated with significant infant mortality and long-term morbidity, and is responsible for more than half of the deaths from congenital anomalies in infancy.[5] Furthermore, CHD is associated with a nine-fold increased risk of intellectual disability.[6]

Timely prenatal detection of CHD therefore has important implications. When a severe defect is detected before 24 weeks of gestation, parents may decide to terminate the pregnancy. When the pregnancy is continued, a prenatal diagnosis of CHD allows time to prepare for the arrival of a sick child and facilitates necessary changes in obstetric and neonatal management. Prenatal diagnosis of CHD was shown to increase survival rates and to decrease long-term morbidity, especially in ductal dependent lesions.[7–13]

Antenatal detection of CHD currently relies on ultrasound scans. In most high-income countries, a standard screening is performed around 20 weeks of gestation, with the assessment of the fetal heart being an important and difficult component. After extensive training, 59.7% of severe CHD and 44.2% of isolated severe CHD were detected antenatally in the Netherlands between 2007 and 2012.[14] After inclusion of the three-vessel view in 2012, the detection rates for transposition of the great arteries and Tetralogy of Fallot were raised from 44% to 82% and 68%, respectively, increasing the overall detection rate.[15] Even though this detection rate exceeds rates reported in most western countries,[16–20] and as has been recently shown for coarctation of the aorta,[21] even in a well-organized screening setting there is still room for further improvement.

Such improvement might come from the addition of the three-vessel-and-trachea view to standard screening protocols,[21] or—as investigated in this paper—non-invasive fetal electrocardiography (ECG). Recent advances enable fetal ECG assessment at gestational ages of about 18 weeks and above[22,23] and in several CHD cases a change in fetal ECG waveform has been shown.[24–29]

In the past, artificial intelligence has been adopted for the classification of CHD from fetal cardiac ultrasound images[30] as well as from fetal heart sounds,[31] and the combination of neonatal ECGs with artificial intelligence has also shown great promise.[32,33] Even though the fetal ECG was shown to be a reliable method for determining the fetal heartrate during labor,[34] little research on the detection of CHD with fetal ECG has been performed because of challenges in both the acquisition and interpretation of the fetal ECG.[35–37] Combining the potential of artificial intelligence and leveraging recent advancements in the measurement of non-invasive fetal ECG, we investigate a novel method for the antenatal detection of CHD.

Key message

Detection of congenital heart disease based on non-invasive fetal electrocardiography shows sensitivity to cases typically missed with conventional ultrasound screening. Detection rates might therefore improve with the addition of electrocardiography-based screening around 20 weeks of gestation.

2 | MATERIAL AND METHODS

2.1 | Participants

Measurements were performed at the Máxima Medical Center, 'Diagnostiek voor U' diagnostic center Eindhoven, Amsterdam Medical Center, Radboud University Medical Center (Radboudumc), Leiden University Medical Center, and Maastricht University Medical Center, all in the Netherlands.

The study population was divided into two groups, a cohort of healthy fetuses and a group of fetuses with possible CHD. Pregnant women 18 years of age or older, with an uneventful pregnancy, carrying a healthy singleton fetus with a gestational age between 18 and 24 weeks were eligible for the healthy cohort. Exclusion criteria were known congenital anomalies and insufficient understanding of the Dutch language. Measurements of the healthy cohort were performed directly before or after the 20-week fetal anomaly scan and after obtaining written informed consent. Three months postnatally, participants received a questionnaire to verify the absence of CHD or other congenital anomalies, as by this time, all children in the Netherlands have undergone a health check-up by a physician. Measurements were excluded when CHD or other congenital anomalies were found postnatally.

Women carrying a fetus with a possible CHD on the fetal anomaly scan were recruited in dedicated centers (Amsterdam Medical Center, Leiden University Medical Center, Maastricht University Medical Center, Radboudumc, Máxima Medical Center), after being referred for an advanced ultrasound examination. After obtaining written informed consent, the ECG measurement was performed directly before or after the advanced ultrasound examination. Participants had to be 18 years of age or older with a gestational age between 18 and 30 weeks. Exclusion criteria were multiple pregnancies and insufficient understanding of the Dutch language. Participants were excluded from further analysis if no CHD was found on postpartum examinations. The CHD diagnosis was confirmed by postnatal echocardiogram/computed tomography scan report/catheterization report/operation report, and when a pregnancy was terminated, by postmortem examination where possible.

CHD measurements were categorized and evaluated for the requirement of an intervention (in the neonatal period, during the first year, or at any stage) by an experienced pediatric cardiologist.

2.2 | Data acquisition and processing

The fetal ECG was recorded with adhesive Ag/AgCl electrodes on the abdomen of the pregnant women lying in a semi-recumbent position. In total, eight electrodes were placed on the abdomen in a fixed configuration as described by Verdurmen et al. for the purpose of the original study.[38]

This provided six bipolar channels of electrophysiological measurements as two electrodes served as a common reference and ground. The duration of the registration was approximately 30 min.

The fetal anomaly scan and advanced ultrasound examination were performed by certified and experienced sonographers.

The electrophysiological signals were digitized and stored at 500 Hz sampling frequency by a prototype fetal monitoring system (Nemo Healthcare BV). After digitization, the acquired signals were processed by PC-based signal processing techniques as previously described by Vullings et al.,[39] Warmerdam et al.,[40,41] and Fotiadou et al.[42] to remove interferences such as the maternal ECG, power-line interference, and electromyographic signals from within the maternal body.

This signal processing was performed on four of the six channels, to allow the use of previously optimized algorithms for the detection of QRS-complexes.[34] For every recording, the same four electrodes were used. If fewer than 25 fetal QRS complexes per minute were detected, the recording was excluded from further analysis due to low signal quality.

The quality of the fetal ECG was enhanced by taking 5 min of consecutive cardiac cycles and splitting the fetal ECG signal into segments containing individual cardiac cycles and subsequently synchronizing segments on the position of the R-peak, after which the median value was taken for each sample. This automated process resulted in the desired per-sample median, i.e. a single ECG segment where temporally uncorrelated noise was suppressed. An example of such a fetal ECG segment is presented in Supporting Information Appendix S1. Measurements lasted roughly 30 min, from which an average of 12 relatively high-quality fetal ECG segments were automatically generated per fetus, using 50% overlap between the 5-min windows. Even though a standardized electrode placement was used, these fetal ECG segments were not corrected for fetal position as the proposed artificial intelligence method was expected to learn to adapt to uncorrected fetal ECGs.

2.3 | Artificial intelligence

The detection of CHD from the fetal ECG segments was based on an artificial neural network. These networks are inspired by biological neural networks, and contrary to programming them for a specific task, these models are trained by adjusting the model weights based on their performance on a set of measurements. These model weights are fixed after training finishes, leaving an algorithm that can be employed in real-world scenarios. In our approach, a model was trained to derive a CHD score, p_{CHD}, for each derived fetal ECG segment. Details of the network and its training can be found in Appendix S1, while an explanation of the network's internal function may be found in Appendix S2 and the PyTorch implementation is given in Appendix S3.

Before training the neural network, the measurements of 122 patients were selected for testing the performance of the network and were removed from the training data set. For this, the healthy and CHD measurements were divided blindly between the training and test data sets. Even though more fetal ECG data are available, the presented model was trained on the data of this single study only. This was done to ensure that all measurements were taken with the same hardware and electrode configurations, preventing the artificially intelligent model from recognizing acquisition setups and consequently making biased CHD predictions based on imbalances between data sets.

To account for the small data set, the training data was used to train 10 different instances of the network. In this approach, the training data were randomly split into 10 groups, ensuring an equivalent fraction of CHD in each group. For each of these model instances, one of the groups was used for validation of the model, while it was trained on the remaining nine groups. These validation measurements were only seen by the model after it was trained.

By using this approach, often referred to as *k*-fold cross-validation, it is possible to get a reliable estimate of the model performance on a relatively small data set, while preventing an over- or underestimation of the model performance as a consequence of the chosen training/validation split for a single instance.[43] This 10-fold cross-validation was carried out without using the 122 measurements of the test set during training.

For each fold, the model performance was evaluated for the accuracy, sensitivity, specificity and positive predictive value (PPV), given by

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN},$$

$$Sensitivity = \frac{TP}{TP + FN},$$

$$Specificity = \frac{TN}{TN + FP},$$

$$PPV = \frac{TP}{TP + FP},$$

with *TP* the number of true positives, *TN* the number of true negatives, *FP* the number of false positives and *FN* the number of false negatives. This evaluation was done on the validation set corresponding to each fold, for which classification was performed on all available ECGs with a threshold at $p_{CHD}=0.5$, meaning the ECG was classified as coming from a CHD case when $p_{CHD} \geq 0.5$ and from a healthy control when $p_{CHD} < 0.5$. The evaluation results for the 10 folds can be found in Appendix S1.

2.4 | CHD detection

After the neural network was trained, it was evaluated on the unseen data of the test population. Combining the p_{CHD} for each of the participants' individual ECG segments in a Bayesian manner gave a single accumulated (i.e. posterior) p_{CHD}. For this, no previous knowledge about the presence of CHD was assumed. This p_{CHD} was updated with each newly available ECG as the clinical measurement progressed, increasing the certainty as more ECG segments were included, and therefore profiting from the length of the measurements.

This accumulated CHD score allowed for the compilation of a receiver-operating-characteristic (ROC) curve for the detection threshold. The performance was then reported by means of an area under the ROC curve (AUC) and the accuracy, sensitivity, specificity, and positive predictive value at the optimal detection threshold were determined. This optimal threshold was calculated as given by the point with maximum Youden's J-statistic, which is the point with maximum distance to random chance.[44] The model's sensitivity was evaluated at a different threshold for the classification of positives as well, moving away from optimal Youden's *J*-statistic towards a specificity of 95%.

Furthermore, tests were performed to evaluate the influence of measurement length and gestational age on the detection performance given this accumulated CHD score.

In addition to the median ECG waveform, the essential QRS-complex detection in our data processing can be used to evaluate the fetal heart rate (FHR) over time. As CHD has been associated with changes in FHR,[28,45] an experiment was performed with the addition of FHR to the network. In this experiment, a second encoder was used for the FHR, and its output was concatenated with the encoded fetal ECG before the fully connected layers.

2.5 | Ethics statement

The study was approved by the Máxima Medical Center institutional review board (NL48535.015.14) on May 8, 2014 and the ECGs used were part of the study described by Verdurmen et al.[38]

3 | RESULTS

In total, 496 measurements were performed; 328 for the healthy group, and 168 measurements for the CHD group, one measurement for each group was lost as the result of data storage problems. After moving two patients from the normal to the CHD group because of a postnatal diagnosis of CHD, and excluding 14 patients who were lost to follow up, a total of 311 measurements from the healthy group were available for further analysis.

From the 168 measurements of the CHD group, four patients were lost to follow up and seven children had a structurally normal heart postnatally. Twelve fetuses were diagnosed with an arrhythmia only and were therefore excluded, and five had to be excluded because of insufficient signal quality. Furthermore, one patient from the originally healthy group, postnatally diagnosed with CHD, had to be excluded as detailed information on the diagnosis was not available. Hence, 140 CHD measurements were left for analysis. The inclusion process and consequent division into train and test sets is depicted in Figure 1.

An overview of patient characteristics can be found in Table 1. Only gestational age during the measurement was significantly different between the two groups. An overview of the incidence of the various types of CHD in this cohort is provided in Table 2, along with ultrasound screening detection rates as found by van Velzen et al.[14] for those groups and algorithm detection rates for all CHD and CHD requiring intervention stratified by the time of the intervention was deemed necessary (as a neonate, in the first year or after a year).

3.1 | Ten-fold performance

Median values for the accuracy, sensitivity, specificity, and positive predictive value on the individual ECGs in the validation set for each of the 10 training folds all exceeded 0.76, and none of the interquartile ranges surpassed 0.14. This 10-fold performance is further specified in Appendix S1.

3.2 | CHD detection performance

Considering the accumulated CHD score assigned to each fetus after evaluating the test data, the ROC curve given in Figure 2 was compiled. The area under this ROC curve was 0.76 and the accuracy, sensitivity, specificity, and positive predictive value at the threshold with maximum Youden's J-static (p_{CHD}=0.0005) were 71%, 63%, 77%, and 71%, respectively.

Table 2 gives an overview of the number of detected CHDs for each group, divided into all CHD cases, cases requiring a neonatal intervention, cases requiring an intervention in the first year, and cases requiring an intervention at any stage. At optimum Youden's J-statistic, the sensitivity was 69% for patients where an intervention was deemed necessary in the first year of life or at a later stage, and the detection rate was 75% for CHDs requiring a neonatal intervention.

Table 3 gives an overview of the sensitivity when operating the proposed algorithm at 95% specificity. As can be seen, overall sensitivity drops to 33.3% for all cases and 40.5% for cases requiring intervention.

3.3 | Effect of measurement length and gestational age

Using the proposed Bayesian updating rule, the performance with respect to measurement length may be evaluated. Figure 3(A) gives

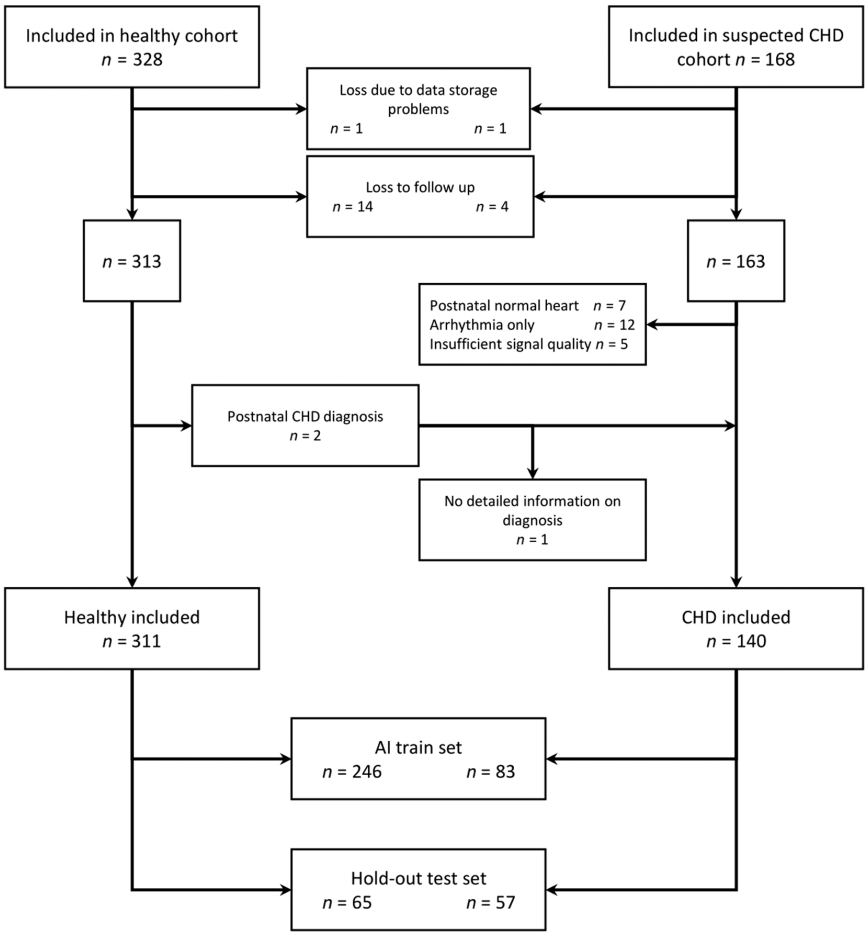

FIGURE 1 Flow diagram of patient inclusions and the consequent division in train and hold-out test sets as used in this paper.

TABLE 1 An overview of patient characteristics.

	Healthy cohort ($n = 311$), mean ± SD	CHD group ($n = 140$), mean ± SD
Age (year)	31.3 ± 4.2	30.6 ± 4.7
BMI (kg/m^2)	24.3 ± 5.2	24.4 ± 4.4
Gestational age (weeks)[a]	20.1 ± 1.1	23.2 ± 3.3
Gravidity	2.0 ± 1.1	2.2 ± 1.5
Parity	0.6 ± 0.7	0.8 ± 0.9

Abbreviations: BMI, body mass index; CHD, congenital heart disease; SD, standard deviation.

[a]$p < 0.05$.

the performance in terms of AUC for different maximum amounts of ECG segments used per individual. Performance increased with an increase in ECG segments used, and the optimal performance was reached when using 14 segments or more, corresponding to a minimum measurement time of 37.5 min. Additionally, Figure 3(B) shows the number of extracted ECG segments per patient for the 122 measurements in the test set.

Figure 4 shows the accuracy of the implemented detection as a function of the gestational age during the measurement. A small negative trend between gestational age and classification accuracy was observed.

3.4 | Use of fetal heart rate

When adding the FHR to the network, an improvement of the 10-fold validation performance could be observed. However, no significant change to the performance on the test data set was found.

4 | DISCUSSION

In this paper, a new method for prenatal screening for CHD was explored. The method can detect CHD at around 20–27 weeks of gestation with a sensitivity and specificity of 63% and 77%, respectively. Most importantly, it can detect 75% and 69% of the fetuses requiring a neonatal intervention or an intervention in the first year, respectively. This performance is independent of the expertise and experience of a sonographer performing the prenatal screening, as is the case with ultrasound-based screening.[14] The presented method uses 5 min of data to provide a single fetal ECG, and is capable of obtaining a fetal ECG in 99.0% of patients.

Our method of using multiple ECG segments per patient increased the detection performance. Even though the performance appears to be optimal for a measurement of at least 37.5 min, the number of ECG segments available per measurement should be considered. The sudden saturation of performance can be explained by only a single measurement of the test set having more than 14 ECG segments.

TABLE 2 An overview of the ultrasound screening detection rates along with the number of samples in the training set and algorithm detection rates for all congenital heart disease (CHD) and CHD requiring an intervention in the neonatal period, during the first year and at any stage in life, stratified per CHD type.

Category	Screening detection[14]	Number in training set	Algorithm detection		CHD requiring an intervention					
					Neonatally		During first year		At any stage	
Septal defects	50.4%	20	8/11	72.7%	0/0	n/a	4/5	80%	5/6	83.3%
Valvular anomalies	32.3%	1	0/1	0.0%	0/0	n/a	0/0	n/a	0/0	n/a
Venous return anomalies	11.1%	0	0/2	0.0%	0/0	n/a	0/0	n/a	0/0	n/a
Aortic arch anomalies	29.7%	10	7/11	63.6%	2/3	66.7%	4/6	66.7%	5/8	62.5%
Conotruncal anomalies	59.8%	31	12/16	75.0%	9/12	75.0%	12/16	75.0%	12/16	75.0%
Hypoplastic right heart syndrome	66.7%	0	1/1	100%	1/1	100%	1/1	100%	1/1	100%
Hypoplastic left heart syndrome	97.6%	3	2/2	100%	2/2	100%	2/2	100%	2/2	100%
Other univentricular heart defects	94.9%	14	3/6	50%	3/4	75.0%	3/6	50.0%	3/6	50.0%
Complex defects with isomerism	93.5%	1	0/3	0.0%	0/1	0.0%	0/2	0.0%	0/2	0.0%
Miscellaneous	20.8%	3	3/4	75%	1/1	100%	1/1	100%	1/1	100%
Total	59.7%	83	36/57	63.2%	18/24	75.0%	27/39	69.2%	29/42	69.0%

Furthermore, the AUC increased from 0.76 to 0.78 when considering only the 101 participants for whom more than eight ECG segments were available, again stressing the relevance of the implemented method of combining individual ECG segments, and the need for research on longer measurements.

There was a small dependency of the classification accuracy on gestational age. As the cardiac intervals are known to change with fetal maturation,[46] this may be influenced by the significant difference in gestational age between the healthy and CHD fetuses. Moreover, this dependency may be influenced by the lower number of available training and/or test samples available in the higher gestational ages.

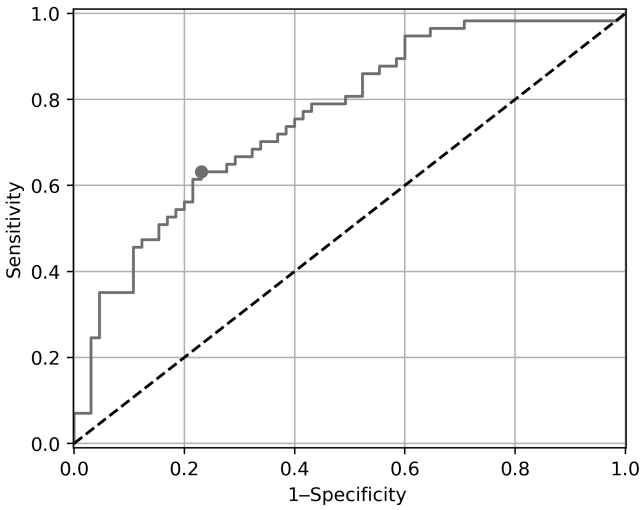

FIGURE 2 Receiver operating characteristics curve for the detection of congenital heart disease considering the combined probability for all available electrocardiographs per patient in the test data. The marked dot gives the point with maximum Youden's J-statistic, corresponding to a decision threshold at 0.05%.

The number of fetuses with CHD in the training set was limited, which may have affected the algorithm's ability to detect specific CHD categories. For instance, no cases from the test data classified as venous return anomaly were detected by the algorithm, which may be explained by the absence of cases with such an anomaly in the training data. However, even though cases of hypoplastic right heart syndrome (HRHS) were lacking in the training data as well, HRHS was successfully detected in the test data. Possibly a persistent left superior caval vein has no effect on the fetal ECG, but HRHS does. Even though Appendix S2 gives some insight into the internal workings of the model, we do not know what the computer sees as discriminant between CHD and non-CHD, making it difficult to identify what factors in the training cohort affect the algorithm's detection ability.

As ECG complexes were combined to enhance the quality of the signals, inter-beat variations in the morphology of the ECG were lost and we could only detect structural electrical abnormalities. With additional signal enhancement methods, we may be able to reduce the number of heartbeats necessary for averaging to the extent where inter-beat variations can become visible and the method could also be used to study arrhythmias.

The sensitivities of the implemented ECG-based detection were highest for hypoplastic left and right heart syndromes (100%), but these diagnoses were sparsely represented in the test cohort. More importantly, the detection rate was good for some CHD that have low detection rates in routine screening, namely conotruncal and aortic arch anomalies, which made up almost half of our test group.

Although the reported detection rates generally outperform those found for sonographic screening, the specificity was 77%.

TABLE 3 Algorithm sensitivity when operating the algorithm at a specificity of 95%. Sensitivities are given for all congenital heart disease (CHD) and CHD requiring intervention at different stages of life, stratified per CHD type.

			CHD requiring an intervention					
Category	**Algorithm detection**		**Neonatally**		**During first year**		**At any stage**	
Septal defects	6/11	54.5%	0/0	n/a	4/5	80%	5/6	83.3%
Valvular anomalies	0/1	0.0%	0/0	n/a	0/0	n/a	0/0	n/a
Venous return anomalies	0/2	0.0%	0/0	n/a	0/0	n/a	0/0	n/a
Aortic arch anomalies	3/11	27.3%	1/3	33.3%	2/6	33.3%	3/8	37.5%
Conotruncal anomalies	6/16	37.5%	4/12	33.3%	6/16	37.5%	6/16	37.5%
Hypoplastic right heart syndrome	0/1	0.0%	0/1	0.0%	0/1	0.0%	0/1	0.0%
Hypoplastic left heart syndrome	2/2	100%	2/2	100%	2/2	100%	2/2	100%
Other univentricular heart defects	1/6	16.7%	1/4	25.0%	1/6	16.7%	1/6	16.7%
Complex defects with isomerism	0/3	0.0%	0/1	0.0%	0/2	0.0%	0/2	0.0%
Miscellaneous	1/4	25.0%	0/1	0.0%	0/1	0.0%	0/1	0.0%
Total	19/57	33.3%	8/24	33.3%	15/39	38.5%	17/42	40.5%

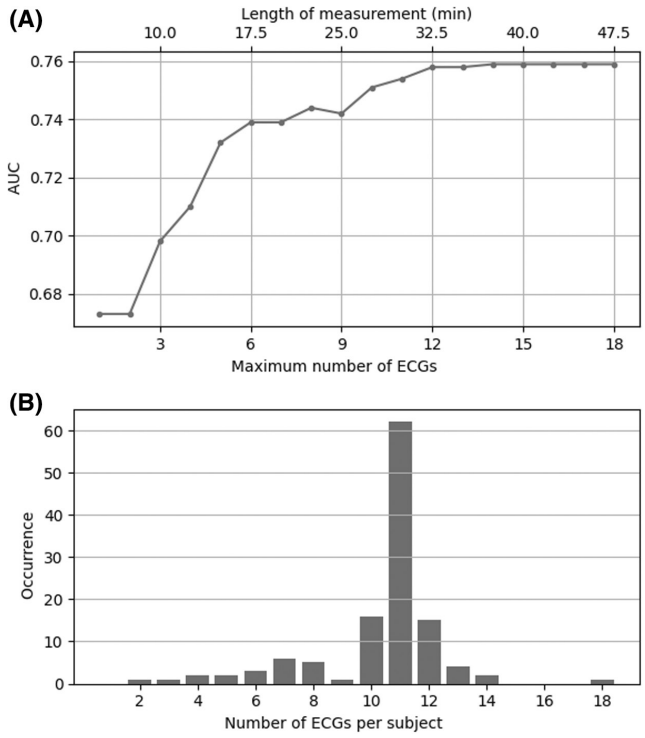

FIGURE 3 (A) The effect of the maximum measurement length on the performance. (B) The number of electrocardiographs (ECGs) extracted per measurement for the 121 measurements in the test data set.

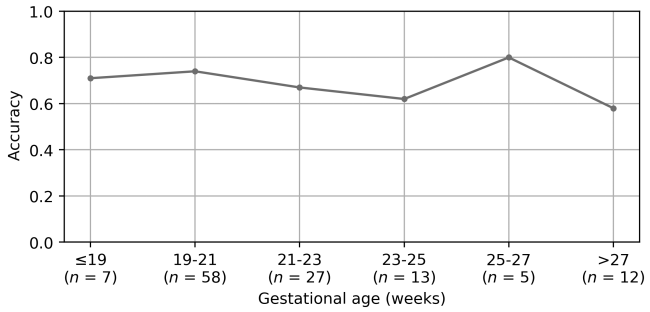

FIGURE 4 Classification accuracy as a function of the gestational age during measurement, the given performance and numbers refer to the test set only, but conform to the distribution as seen in the training data.

Even though a lower specificity might be acceptable because the implications of missing an important diagnosis as opposed to a false positive are much worse, the number of false-positive cases and their implications may be reduced by using a different threshold for the classification of positives. Despite the model's sensitivity dropping to 33% for a specificity of 95%, the algorithm still outperforms sonographic screening for the cases with *septal defects* and *aortic arch anomalies* when operated in this way. These cases illustrate how an ECG, based on a different physical principle, shows sensitivity to different types of CHDs. In future work, we therefore envision the exploration of ECG-based screening as a tool complementary to the customary ultrasound screening for CHD. As the current data set does not allow the evaluation of such a combined screening, this work should be considered as a feasibility study, requiring further research.

When considering the use of ECG-based screening complementary to conventional ultrasound screening, the duration of this screening would increase by more than 30 minutes. However, these measurements do not require the presence of a caregiver and are therefore expected to minimally increase workload. With the minimal instruction that is required for these measurements and recent attention for remote electrophysiological monitoring,[47,48] it may even be possible for patients to perform them at home.

Our method achieved higher detection rates than those reported for standard ultrasound screening,[14,16–20,49] but the data set was biased towards more severe CHD. The detection of severe CHD cases is expected to have the highest impact on morbidity and mortality because we can set the required medical care in motion to offer that individual the best perinatal care. Should we miss a small ventricular septal defect that will close spontaneously during the rest of the pregnancy or first year of life, this will not affect the neonate detrimentally and the baby can be born without any fear of problems from the ventricular septal defect during the birth or transition.

Furthermore, it should be mentioned that anatomical malformations might not necessarily be accompanied by electrical abnormalities in the fetus, but subtle abnormalities might be accompanied by relatively large electrical abnormalities that can easily be detected.

Future work should aim at improving the signal processing chain to reduce or omit the need for combining multiple heartbeats, potentially further increasing the performance and allowing for the additional analysis of arrhythmias.

Apart from taking the median of multiple heart beats over time to eliminate temporally uncorrelated noise, the vectorcardiogram might be used to correct for fetal orientation and movement during the measurement.[50] This enhancement of the signal quality by using this three-dimensional representation of the ECG could improve the model performance and reduce the measurement time required to reach optimal performance, both increasing the method's potential. Moreover, the vectorcardiogram might be used to simulate random rotations, which would help to prevent our network from learning to associate fetal position with CHD, which may arise from the currently small data set. Similarly, the detection performance might be improved by incorporating the currently omitted two channels. However, an extensive data set is needed for the optimization of algorithms for removal of interferences and detection of QRS-complexes when using more than four channels. The vectorcardiogram approach works independently of the number of channels and hence is favorable due to being directly applicable to the six-lead measurement.[51]

Furthermore, some CHD manifest themselves by alterations in the orientation of the electrical heart axis. For most of the recordings in our data set, information on the fetal orientation was available and the orientation of the electrical heart axis could be calculated.[36,52] This would yield additional information that could increase the CHD

detection rates and could be the subject of further investigations. Similarly, further research on the possibilities of FHR to improve ECG-based detection should be conducted on a more extensive data set or by including longer FHR fragments.

5 | CONCLUSION

This feasibility study shows that detection rates of CHD might improve by extending the standard ultrasound-based screening with fetal electrocardiography, which is achievable for 99.0% of patients. However, more research is required to improve performance and to determine the benefits to clinical practice.

AUTHOR CONTRIBUTIONS

IdV conducted this research and wrote the manuscript; RV and SC participated in drafting the study design and manuscript. SC classified participants based on pathology type and severeness. RV, JvL, and MvdH supervised this research. All authors revised and approved the manuscript.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Drs. Aimee van Dobben-Rodenburg from 'Diagnostiek voor U' and to N. Eijsvoogel, D. Aben, M. Sengers, J. Drinkwaard, C. van den Oord, M. van Wierst, O. Hulsenboom, M.C. Haak, C. Lempersz, L. Noben, and L. Cornelissen for their role in the data collection.

FUNDING INFORMATION

This research was partly funded by ZonMW as part of the Topspecialistische Zorg en Onderzoek (TZO) project "BIRTHSCREEN" (project number 10070022010001).

CONFLICT OF INTEREST STATEMENT

RV is one of the founders and shareholder of Nemo Healthcare BV, The Netherlands. The remaining authors report no conflict of interest.

ORCID

Ivar R. de Vries https://orcid.org/0000-0002-0579-4220
Judith O. E. H. van Laar https://orcid.org/0000-0001-5350-3267
Marieke B. van der Hout-van der Jagt https://orcid.org/0000-0002-5805-5333
Sally-Ann B. Clur https://orcid.org/0000-0001-5276-1475
Rik Vullings https://orcid.org/0000-0002-2392-6098

REFERENCES

1. Dolk H, Loane M, Garne E. European surveillance of congenital anomalies working G. *Congenit Hear Defects Eur Preval Perinat Mortal*. 2000;123:841-849.
2. Hoffman JIE, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol*. 2002;39:1890-1900.
3. Nelle M, Raio L, Pavlovic M, et al. Prenatal diagnosis and treatment planning of congenital heart defects—possibilities and limits. *World J Pediatr*. 2009;5:18-22.
4. Van Der Linde D, Konings EEM, Slager MA, et al. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol*. 2011;58:2241-2247.
5. Young ID, Clarke M. Lethal malformations and perinatal mortality: a 10 year review with comparison of ethnic differences. *Br Med J (Clin Res ed)*. 1987;295:89-91.
6. Razzaghi H, Oster M, Reefhuis J. Long-term outcomes in children with congenital heart disease: National Health Interview Survey. *J Pediatr*. 2015;166:119-124.e1.
7. Mahle WT, Clancy RR, McGaurn SP, et al. Impact of prenatal diagnosis on survival and early neurologic morbidity in neonates with the hypoplastic left heart syndrome. *Pediatrics*. 2001;107:1277-1282.
8. Tworetzky W, McElhinney DB, Reddy VM, et al. Improved surgical outcome after fetal diagnosis of hypoplastic left heart syndrome. *Circulation*. 2001;103:1269-1273.
9. Franklin O, Burch M, Manning N, Sleeman K, Gould S, Archer N. Prenatal diagnosis of coarctation of the aorta improves survival and reduces morbidity. *Heart*. 2002;87:67-69.
10. Kaguelidou F, Fermont L, Boudjemline Y, le Bidois J, Batisse A, Bonnet D. Foetal echocardiographic assessment of tetralogy of Fallot and post-natal outcome. *Eur Heart J*. 2008;29:1432-1438.
11. Trines J, Fruitman D, Zuo KJ, Smallhorn JF, Hornberger LK, Mackie AS. Effectiveness of prenatal screening for congenital heart disease: assessment in a jurisdiction with universal access to health care. *Can J Cardiol*. 2013;29:879-885.
12. Hunter LE, Simpson JM. Prenatal screening for structural congenital heart disease. *Nat Rev Cardiol*. 2014;11:323-334.
13. Van Velzen CL, Haak MC, Reijnders G, et al. Prenatal detection of transposition of the great arteries reduces mortality and morbidity. *Ultrasound Obstet Gynecol*. 2015;45:320-325.
14. Van Velzen CL, Clur SAB, Rijlaarsdam MEB, et al. Prenatal detection of congenital heart disease - results of a national screening programme. *BJOG*. 2016;123:400-407.
15. Everwijn SM, Van der Palen R, van Velzen C, et al. OP20.07: a prenatal detection rate of 85% is achieved for transposition of the great arteries after introduction of the three-vessel view. *Ultrasound Obstet Gynecol*. 2016;48:116-117.
16. Levy DJ, Pretorius DH, Rothman A, et al. Improved prenatal detection of congenital heart disease in an integrated health care system. *Pediatr Cardiol*. 2013;34:670-679.
17. Buskens E, Grobbee DE, Frohn-Mulder IME, et al. Efficacy of routine fetal ultrasound screening for congenital heart disease in Normal pregnancy. *Circulation*. 1996;94:67-72.
18. Sharland GK, Allan LD. Screening for congenital heart disease prenatally. Results of a 2½-year study in the south East Thames region. *BJOG*. 1992;99:220-225.
19. Stümpflen I, Stümpflen A, Wimmer M, Bernaschek G. Effect of detailed fetal echocardiography as part of routine prenatal ultrasonographic screening on detection of congenital heart disease. *Lancet*. 1996;348(9031):854-857.
20. Galindo A, Herraiz I, Escribano D, Lora D, Melchor JC, de la Cruz J. Prenatal detection of congenital heart defects: a survey on clinical practice in Spain. *Fetal Diagn Ther*. 2011;29:287-295.
21. Zwanenburg F, Ten Harkel ADJ, Snoep MC, et al. Prenatal detection of aortic coarctation in a well-organized screening setting: are we there yet? *Prenat Diagn*. 2023;43:620-628.
22. Behar J, Andreotti F, Zaunseder S, Oster J, Clifford GD. A practical guide to non-invasive foetal electrocardiogram extraction and analysis. *Physiol Meas*. 2016;37:R1-R35.
23. Vullings R. *Non-Invasive Fetal Electrocardiogram: Analysis and Interpretation*. Eindhoven University of Technology; 2010.
24. Verdurmen KMJ, Eijsvoogel NB, Lempersz C, et al. A systematic review of prenatal screening for congenital heart disease by fetal electrocardiography. *Int J Gynecol Obstet*. 2016;135:129-134.

25. Hamilton LA, Fisher E, Horn C, et al. A new prenatal cardiac diagnostic device for congenital heart disease. *Obstet Gynecol*. 1977;50:491-494.
26. Brambati B, Bonsignore L. The significance of indirect electrocardiography in fetal cardiac arrhythmias. *Eur J Obstet Gynecol Reprod Biol*. 1983;14:371-373.
27. Velayo C, Sato N, Ito T, et al. Understanding congenital heart defects through abdominal fetal electrocardiography: case reports and clinical implications. *J Obstet Gynaecol Res*. 2011;37:428-435.
28. Siddiqui S, Wilpers A, Myers M, Nugent JD, Fifer WP, Williams IA. Autonomic regulation in fetuses with congenital heart disease. *Early Hum Dev*. 2015;91:195-198.
29. Yilmaz B, Narayan HK, Wilpers A, Wiess C, Fifer WP, Williams IA. Electrocardiographic intervals in foetuses with CHD. *Cardiol Young*. 2016;26:84-89.
30. Day TG, Kainz B, Hajnal J, Razavi R, Simpson JM. Artificial intelligence, fetal echocardiography, and congenital heart disease. *Prenat Diagn*. 2021;41:733-742.
31. Hoodbhoy Z, Jiwani U, Sattar S, Salam R, Hasan B, Das JK. Diagnostic accuracy of machine learning models to identify congenital heart disease: a meta-analysis. *Front Artif Intell*. 2021;4:708365.
32. Yang S, Yamauchi K, Nonokawa M, Ikeda M. Use of an artificial neural network to differentiate between ECGs with IRBBB patterns of atrial septal defect and healthy subjects. *Med Inform Internet Med*. 2002;27:49-58.
33. Someshwaran G, Sarada V. A Research Review on Fetal Heart Disease Detection Techniques. Proc - 2022 6th Int Conf Intell Comput Control Syst ICICCS 2022 2022;(Icicccs):1674–1681. doi:10.1109/ICICCS53718.2022.9788226
34. Vullings R, van Laar JOEH. Non-invasive fetal electrocardiography for intrapartum cardiotocography. *Front Pediatr*. 2020;8:599049.
35. Lempersz C, Van Laar JOEH, Clur SAB, et al. The standardized 12-lead fetal electrocardiogram of the healthy fetus in mid-pregnancy: a cross-sectional study. *PLoS One*. 2020;15:1-12.
36. Lempersz C, Noben L, Clur SAB, et al. The electrical heart axis of the fetus between 18 and 24 weeks of gestation: a cohort study. *PLoS one*. 2021;16:1-11.
37. Noben L, Lempersz C, van den Heuvel E, et al. The electrical heart axis in fetuses with congenital heart disease, measured with noninvasive fetal electrocardiography. *PLoS One*. 2022;17:1-15.
38. Verdurmen KMJ, Lempersz C, Vullings R, et al. Normal ranges for fetal electrocardiogram values for the healthy fetus of 18-24 weeks of gestation: a prospective cohort study. *BMC Pregnancy Childbirth*. 2016;16:1-7.
39. Vullings R, Peters CHL, Sluijter RJ, Mischi M, Oei SG, Bergmans JWM. Dynamic segmentation and linear prediction for maternal ECG removal in antenatal abdominal recordings. *Physiol Meas*. 2009;30:291-307.
40. Warmerdam GJJ, Vullings R, Schmitt L, van Laar JOEH, Bergmans JWM. A fixed-lag Kalman smoother to filter power line interference in electrocardiogram recordings. *IEEE Trans Biomed Eng*. 2016;64:1852-1861.
41. Warmerdam GJJ, Vullings R, Schmitt L, van Laar JOEH, Bergmans JWM. Hierarchical probabilistic framework for fetal R-peak detection, using ECG waveform and heart rate information. *IEEE Trans Signal Process*. 2018;66:4388-4397.
42. Fotiadou E, van Laar JOEH, Oei SG, Vullings R. Enhancement of low-quality fetal electrocardiogram based on time-sequenced adaptive filtering. *Med Biol Eng Comput*. 2018;56:2313-2323.
43. Wong TT, Yeh PY. Reliable accuracy estimates from k-fold cross validation. *IEEE Trans Knowl Data Eng*. 2019;32:1586-1594.
44. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32-35.
45. Ueda K, Ikeda T, Iwanaga N, et al. Intrapartum fetal heart rate monitoring in cases of congenital heart disease. *Am J Obstet Gynecol*. 2009;201:64.e1-64.e6.
46. Smith V, Arunthavanathan S, Nair A, Ansermet D, da Silva CF, Wallace EM. A systematic review of cardiac time intervals utilising non-invasive fetal electrocardiogram in normal fetuses. *BMC Pregnancy Childbirth*. 2018;18:370.
47. Zizzo AR, Hvidman L, Salvig JD, Holst L, Kyng M, Petersen OB. Home management by remote self-monitoring in intermediate- and high-risk pregnancies: a retrospective study of 400 consecutive women. *Acta Obstet Gynecol Scand*. 2022;101:135-144.
48. Manna S, Sriraam N, Pandian PS. Prototype of Home Based Multi-Channel Wearable Wireless Fetal ECG Monitoring System. Proc CONECCT 2021 7th IEEE Int Conf Electron Comput Commun Technol; 0–3. 2021. doi:10.1109/CONECCT52877.2021.9622705
49. Marek J, Tomek V, Škovránek J, et al. Prenatal ultrasound screening of congenital heart disease in an unselected national population: a 21-year experience. *Heart*. 2011;97:124-130.
50. Vullings R. Fetal electrocardiography and deep learning for prenatal detection of congenital heart disease. *Comput Cardiol Conf*. 2019;45:1-4.
51. Vullings R, Mischi M. Vectorcardiographic loop alignment for fetal movement detection using the expectation-maximization algorithm and support vector machines. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*. 2013;2013:2915-2918. doi:10.1109/EMBC.2013.6610150
52. Verdurmen KMJ, Hulsenboom ADJ, van Laar JOEH, Wijn PFF, Vullings R, Oei SG. Orientation of the electrical heart axis in midterm pregnancy. *Eur J Obstet Gynecol Reprod Biol*. 2016;207:243-246.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: de Vries IR, van Laar JOEH, van der Hout-van der Jagt MB, Clur S-A, Vullings R. Fetal electrocardiography and artificial intelligence for prenatal detection of congenital heart disease. *Acta Obstet Gynecol Scand*. 2023;102:1511-1520. doi:10.1111/aogs.14623