Physiological
Measurement

PAPER • **OPEN ACCESS**

Low-complexity fetal heart rate monitoring from carbon-based single-channel dry electrodes maternal electrocardiogram

To cite this article: S Likitalo *et al* 2026 *Physiol. Meas.* **47** 015006

IOP Publishing Physiol. Meas. 47 (2026) 015006 https://doi.org/10.1088/1361-6579/ae3365

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

PAPER

Low-complexity fetal heart rate monitoring from carbon-based single-channel dry electrodes maternal electrocardiogram

S Likitalo[1,*], A Anzanpour[2], A Axelin[1], T Jaako[3] and P Celka[4]

[1] Department of Nursing Science, University of Turku, Kiinamyllynkatu 10, 20520 Turku, Finland
[2] Department of Computing, University of Turku, Yliopistonmäki, Vesilinnantie 5, 20500 Turku, Finland
[3] Polar Electro Oy, Professorintie 5, 90440 Kempele, Finland
[4] Polar Electro Europe AG, Av. D.-Jeanrichard 2a, 2114 Fleurier, Switzerland
[*] Author to whom any correspondence should be addressed.

E-mail: smliki@utu.fi

Keywords: heart rate, fetus, electrocardiogram, machine learning, Viterbi, Polar H10

OPEN ACCESS

RECEIVED
3 October 2025

REVISED
8 December 2025

ACCEPTED FOR PUBLICATION
5 January 2026

PUBLISHED
14 January 2026

Abstract

Objective. Fetal and maternal health during pregnancy can be monitored with sensors such as Doppler or scalp fetal ECG. This study focuses on single-channel dry electrode maternal abdominal ECG (*aECG*) to extract fetal heart rate (*fHR*) using a low-complexity algorithm suitable for low-power wearables. *Approach*. A hybrid model combining machine learning, QRS masking, and data fusion was trained on two PhysioNet databases and synthetically generated *aECG*. Model selection employed the Akaike criterion with data balancing and random sampling. *Main results*. The algorithm was tested on 80 recordings from the Computer in Cardiology Challenge 2013 (CCC) and the abdominal and direct fetal database (ADFD), augmented with 100 synthetic *aECG*. Performance for fetal QRS detection reached *Precision* = 97.2(82.2)%, *Specificity* = 99.8(93.8)%, and *Sensitivity* = 97.4(93.9)% on ADFD and CCC, respectively. Clinical validation used the Polar Electro Oy H10 dry-electrode device at the Maternity Hospital of Southwest Finland. Four subjects (gestational age 39.8 ± 1.3 weeks) were analyzed, with seven discarded. For *fHR*, the mean absolute percentage error was 1.9 ± 1.0%, Availability 79.6 ± 3.9%, and coverage probability $CP_5 = 76.2$%, $CP_{10} = 87.5$%. *Significance*. These results demonstrate the feasibility of *fHR* monitoring from dry-electrode *aECG* tailored for low-power wearables. Signal quality in clinical subjects matched the lowest PhysioNet cases, confirming robustness under low signal-to-noise conditions.

1. Introduction

Between 2015 and 2019, an estimated 275 million expectant mothers were reported annually worldwide (Bearak *et al* 2020). Pregnancy is a critical period in a woman's life, marked not only by profound physiological changes but also by potential health risks for both the mother and the developing fetus. One of the most important indicators of fetal well-being is the fetal heart rate (*fHR*), which reflects the fetus's oxygenation status and overall health.

Careful monitoring of maternal and fetal parameters during pregnancy is therefore essential to detect complications early, guide timely interventions, and improve perinatal outcomes (Hernandez Engelhart *et al* 2023).

Late and variable decelerations in *fHR* may indicate fetal distress, most commonly due to hypoxic stress, which can result in fetal brain injury (Ross 2011, Chandraharan *et al* 2023). The primary purpose of electronic fetal monitoring is to detect such hypoxic events early, thereby reducing the risk of neurological injury.

To avoid undesirable fetal outcomes, identifying changes in *fHR* is essential for timely medical interventions (Hernandez Engelhart *et al* 2023). Similarly, maternal health factors-including blood pressure, heart rate, and underlying conditions-can directly influence fetal well-being.

In current pregnancy care, *fHR* monitoring is based on measurements conducted in health care setting. In Finland, for example, *fHR* is measured during all pregnancy follow-up visits, both in maternity clinic and hospital (Finnish Institute for Health and Welfare 2013). As the measurements are taken at the health care visits, they can only be used to assess the fetal condition at those certain time points. If healthcare professional, based on these measurements, identifies a need for more frequent fetal monitoring to ensure the well-being of the fetus, it usually results in multiple visits to healthcare or even hospitalization.

Nevertheless, it is well known, that being able to stay in home environment is beneficial for pregnant woman's overall well-being (Zizzo *et al* 2022), which supports the need to develop remote monitoring systems. Enabling monitoring possibilities outside health care setting could provide pregnant women an opportunity to stay at their familiar surroundings and still feel safe while being remotely monitored by the health care professionals (O'Brien *et al* 2013). As currently used pregnancy monitoring technologies are not ideal or usable outside healthcare, it is important to develop new reliable technologies enabling remote monitoring also in home environment.

Monitoring health in home environment is challenging for several reasons: 1) the device setup must be as simple and comfortable as possible yet registering the required signals, 2) noise and motion artifacts must be dealt appropriately to capture meaningful metrics (Matonia *et al* 2020), 3) the recording and analysis of the data must be as automatic as possible requiring minimum user intervention, and 4) the user interface, either on the device or a mobile unit, must be simple and informative in order to avoid extra stress with inaccurate metrics (Serrano *et al* 2023). Especially increased stress is identified as a risk in pregnancy and fetal monitoring (Walter *et al* 2022). Furthermore, the integration of technology into remote health care monitoring necessitates a system capable of facilitating seamless and real-time data transmission to health care professionals (Tijus *et al* 2024).

This work presents a proof-of-concept of using a commercial electrocardiogram sensor (Polar Electro Oy, H10) to monitor fetal and pregnant woman's heart rate in home environment. The sensor is made of a skin compatible dry carbon-based thermoplastic polyurethane material with long elastance durability The device has been tested and validated with gold standard clinical ECG systems (Gilgen-Ammann *et al* 2019, Skála *et al* 2022).

In health care setting, current technologies to measure *fHR* are mostly based on cardiotocography using Doppler ultrasound (DU) sensors (Ponsiglione *et al* 2021), requiring health care professionals' input in conducting the measurements and expertise on finding the optimal placement for the sensors. In certain complicated situations during the labor, the *fHR* can also be monitored using fetus scalp electrocardiogram (ECG), but that is an invasive procedure and requires attaching the ECG electrode to the fetus' head (Neilson 2015). Recent development in *fHR* monitoring also includes textile-based solutions (Ahmed *et al* 2024).

To promote the adoption of technology and remote health care services, pregnant women must be equipped to perform the necessary measurements independently within the home setting. In that case, the optimal means for *fHR* monitoring could be through pregnant woman's abdominal ECG (*aECG*). Even though some pregnant women are already using commercial home DU devices, it is well known that DU is sensitive to signal loss related to, for example, pregnant woman's or fetal movements, high BMI of the pregnant woman or irregular *fHR* (Lempersz *et al* 2020). These can also lead to misinterpretation, which could be tolerated by using ECG technology (Liu *et al* 2023). The use of ECG technology to measure fetal well-being could also enable options for continuous and real-time monitoring (Jezewski *et al* 2017, Yuan *et al* 2019). Therefore, the most promising alternative solution is to use an abdominal belt with ECG sensors (Chen *et al* 2025, Li *et al* 2025).

Advances in monitoring technologies, ranging from traditional DU to wearable sensors and Artificial Intelligence-driven analysis, have significantly improved the ability to continuously and non-invasively assess both maternal and fetal health. Previous literature presents versatile options for non-invasive fetal ECG monitoring (Khandoker *et al* 2018, Nichting *et al* 2021). However, there are still challenges in accurately detecting and analyzing fetal cardiac signals. The most significant challenge is related to the signal-to-noise ratio (SNR). As the fetal ECG signal is weaker than the pregnant woman's ECG, it poses challenges to signal separation and extraction (Wang and Zhao 2022), requiring advanced signal processing, signal quality estimation (Pimentel *et al* 2014, Rahman *et al* 2022, Shi *et al* 2023, Zhou *et al* 2025), and filtering techniques (Barnova *et al* 2021). Even though utilization of artificial intelligence could improve the extraction and analysis of fetal ECG signals, it also requires robust and accurate algorithms (Lee and Lee 2022).

Signal processing based *fHR* estimation from pregnant woman's *aECG* includes signal processing based approaches such as singular value decomposition, independent component analysis, adaptive noise cancellation (Andreotti *et al* 2017, Jamshidian-Tehrani and Sameni 2018, Kahankova *et al* 2019, Wang

and Zhao 2022, Yerande *et al* 2022, Darsana and Kumar 2023, Lampros *et al* 2023) and template matching (Jaros *et al* 2024) to machine and deep learning (Silva *et al* 2012, Lee and Lee 2022, Abel *et al* 2023, Darmawahyuni *et al* 2023, Wahbah *et al* 2024, Zhou *et al* 2024, 2025, Chen *et al* 2025, Li *et al* 2025). Hybrid techniques using classic signal processing and machine learning were recently developed (Jaros *et al* 2019, Samuel and Hota 2024, Ziani 2024). We approached the solution using such an hybrid technique based on adjustable pregnant woman QRS template removal, machine learning with model selection and Viterbi-based fusion. This hybrid approach allows us to integrate the best of both techniques.

The paper starts with the Method section 2, where we describe the approach to extract the fetus and pregnant woman heart rates, datasets and recording protocols. Signal processing and heart rate estimation techniques are presented in sections 3 and 4. Section 5 presents the statistical analysis of the heart rate estimations on the different datasets. Section 6 elaborates on the results and their interpretations, and section 7 draws the conclusions.

2. Methods

2.1. Approach

The *aECG* is essentially composed of three signals: 1) the pregnant woman ECG (*mECG*), 2) the fetus ECG (*fECG*), and, 3) a noise term *Noise*. The noise component is therefore everything in the *aECG* which is neither the expectant mother's nor fetal ECG, and may include pregnant woman motion (including breathing) and muscle electromyographic artifacts, as well as electromyographic activities from uterine contractions, sensor intrinsic noise, sensor-to-skin contact, power line interferences. The noise component is especially important when using dry electrodes (Searle and Kirkup 2000, Galli *et al* 2021, Joutsen *et al* 2024, Kumar *et al* 2025). A linear combination of these signals is assumed in this work: $aECG = mECG + fECG + Noise$. The *mECG* and *fECG* are composed of the pregnant woman (*mQRS*) and fetus QRS complexes (*fQRS*).

The time difference between the Q and S waves extrema of the QRS complexes are called the QRS zones as shown in figure 5(a). These zones are different for *mQRS* and *fQRS* and are taken into account in our algorithm, especially during the *mQRS* masking in Branch 1 shown in figure 1. The precise estimation of the peak *R*-wave location for each QRS complex is then performed by searching for the extrema in this QRS zone[5].

The algorithm is made of two complementary Branches and a third one combining the two as shown in figure 1. The *aECG* is processed using a sliding window of 5 s. The *aECG* is first cleaned from artifacts, noise is reduced, and a signal quality index $aECG_{\mathrm{SQI}}$ is computed.

In Branch 1, after some preprocessing as explained in section 3, a peak *R*-wave detection technique based on the Pan-Tompkins (PT) algorithm is used to detects the $mQRS_{\mathrm{PT}}$. A masking technique is then performed on the $mQRS_{\mathrm{PT}}$ complexes and the $fQRS_{\mathrm{PT}}$ are detected using a similar procedure.

In Branch 2, a classifier (C) is trained to recognize three classes: a) the *mQRS*, *fQRS*, and a noise component *Noise*. The outputs of the classifier are labeled: $mQRS_{\mathrm{C}}$, $fQRS_{\mathrm{C}}$, and $Noise_{\mathrm{C}}$. Two *fQRS* quality indices $mfSQI_{\mathrm{PT}}$ from Branch 1 and $mfSQI_{\mathrm{C}}$ from Branch 2 are simultaneously computed. The two quality indices are defined as the average ratio of the maximum amplitude of the *fQRS* and the maximum amplitude of *mQRS* on a segment of 5 s. The choice of 5 s is based on the following constraints: 1) the real-time implementation which limits the number samples per buffer, 2) the minimum mother's heart rate which is typically around 60 bpm. To estimate a heart rate at 60 bpm we need at least 4 heartbeats which amount to 4 s. Furthermore, to account for possible arrhythmias, we used a window of 5 s which is small enough to capture HR variations, and 3) the *fHR* can accelerate and decelerate in a short period of time. The choice of 5 s allows to capture these physiologically relevant events (Pawł owski *et al* 2025). The lower *mfSQI*, the more hidden is the *fQRS*. The *mfSQI* is also expressed in decibels and labeled $mfSQI_{\mathrm{SNR}}$.

In Branch 3, a Viterbi algorithm produces the final *fHR* (Souriau *et al* 2023) from Branch 1 and 2. Sections 3, 4, appendices C and B provide more details on these processing techniques.

2.2. Databases

Physionet

Two databases from Physionet (Goldberger *et al* 2000), namely the Computer in Cardiology 2013 challenge (hereafter Computer in Cardiology Challenge (CCC) (Clifford *et al* 2014), sampling rate 1000 Hz) and the Abdominal and Direct Fetal ECG Database (hereafter abdominal and direct fetal database

[5] We will further use interchangeably the terms $m(f)QRS$ and $m(f)R$ when this distinction is not strictly necessary.

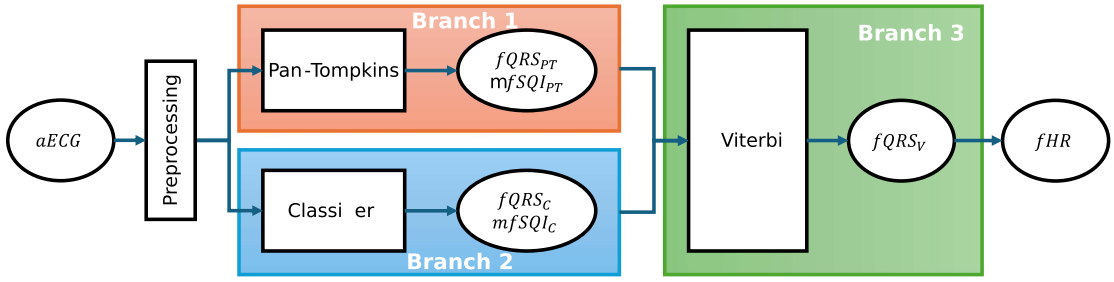

Figure 1. Overview of the algorithm architecture focusing on the fetal QRS detection.

Table 1. ECG model parameters.

Range	*mHR* (bpm)	*fHR* (bpm)	*aECG* (dB)	*mfQRS* (dB)
Min	60	80	10	−60
Max	140	240	20	−10

(ADFD) (Jezewski *et al* 2012), sampling rate 1000 Hz) totaling 80 recordings. All ECG signals were downsampled to 200 Hz.

The ADFD database consists of five different women in labor, between 38 and 41 weeks of gestation. Each recording lasted 5 min each and were acquired on four ECG derivations (channels). The reference fetus peak *R*-wave locations $fR_{\text{ref}}(k)$ (annotations) were extracted from scalp ECG by a group of cardiologists. The recordings were generally of excellent quality, even allowing the fetus's heart rate variability analysis and pregnant woman-fetus heart beat-to-beat coupling.

The CCC database, containing 75 recordings (extended set-a) and lasting 60 s each, was also used in our algorithm development. Each record was acquired using four ECG derivations. No information was found for the gestational age. We noted that the average quality of the *mECG* and *fECG* signals was lower than that of the ADFD ones. The $fR_{\text{ref}}(k)$ were obtained from *crowd-sourcing using a mixture of experts, volunteers, and algorithms* (text extracted from the Physionet CCC website https://physionet.org/content/challenge-2013/1.0.0/).

ECG synthetic model

A synthetic *aECG* generator (SYN) (Sameni *et al* 2007, Andreotti *et al* 2016, Keenan *et al* 2018) containing a linear mixture of *mECG* and *fECG* was used to produce 100 *aECG* signals of 30 s duration. We chose 100 simulations to approximately match the Physionet databases. The idea behind using a digital ECG simulator is: 1) to easily generate a large number of signals, 2) to have highly configurable model parameters, and 3) to simulate ECG electrode placements and various SNRs ($aECG_{\text{SNR}}$). We generated a noise signal that challenged the *fECG* in terms of amplitude to make this simulation even more realistic.

The model simulates the P-QRS-T complexes *et al* (Sameni *et al* 2007) and modified to include fetal ECG by Andreotti *et al* (2016) based on a nonlinear dynamical differential equation. The model can simulate a multichannel ECG with essential parameters influencing the *aECG* shape such as: 1) fetus position and movement, 2) the ECG electrode positions on the abdomen, 3) $aECG_{\text{SNR}}$ and the $mfQRS_{\text{SNR}}$ in units of decibels, 4) pregnant woman heart rate (HR) and variability (HRV): *mHR* and *fHR* in units of beats per minutes (bpm),and mother *mHRV* expressed as the Low Frequency to High Frequency ratio (LF/HF) (Shaffer and Ginsberg 2017), 5) pregnant woman and fetus breathing rate *mBR* and *fBR* in units of cycle per minute (cpm), and 6) uterine contractions.

We simulated the *aECG* with fixed $mBR = 15$ cpm and $fBR = 35$ cpm, and $LF/HF = 0.6$. We did not include uterine contractions. We used randomly (Gaussian distribution) chosen parameters in the ranges described in table 1.

The simulated movement of the fetus was chosen to be helicoidal, which produced different $mfQRS_{\text{SNR}}$ as the fetus's heart dipole projection on the electrodes changes in amplitude (Behar *et al* 2014). In the model of Behar *et al*, there is a possibility to include fetus heart rate accelerations and decelerations. We did not include this feature in our simulations. We downsampled the simulated ECG from 1000 Hz to 200 Hz for consistency with the other databases. The simulator thus produced maternal and fetus QRS complexes *mQRS* and *fQRS*, which are further used for training and testing the classifier in Branch 2 (see figure 1).

Table 2. Subjects' information.

ID	GA (weeks + days)	H10 location	Notes
02	41 + 5	Under the navel	Fetus moves a lot in Phase 1
03	37 + 2	Above the navel	Pregnant woman had light fever and showed agitation
04	39 + 0	Above the navel	Pregnant woman agitated. A lot of amniotic liquid
05	40 + 3	Under the navel	Phase 1 stopped due to contractions
06	39 + 6	Above the navel	—
07	31 + 3	Under the navel	—
08	37 + 6	Under the navel	Pregnant woman had a mall belly
09	38 + 6	Under the navel	Pregnant woman half sitting on the back
10	38 + 0	Under the navel	Obese pregnant woman half sitting on one side
11	40 + 0	Under the navel	—
12	40 + 1	Under the navel	—

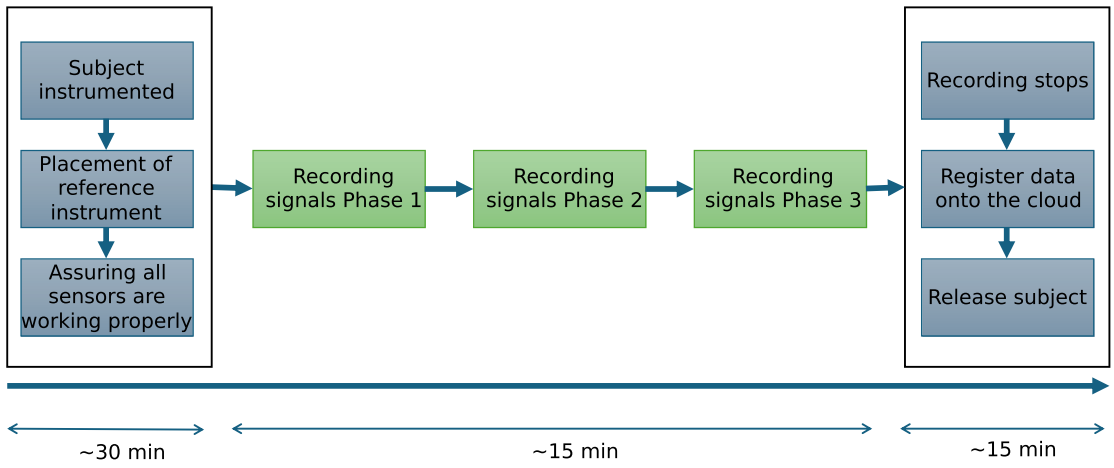

Figure 2. Recording protocol.

Clinical Data

The clinical data were collected in hospital environment at the Pregnancy follow-up ward of the Maternity hospital of the Wellbeing services county of Southwest Finland between February and April 2025. The participants were pregnant women with: 1) minimum 29 weeks of pregnancy, 2) being single pregnancy, 3) age superior to 18 years) the ability to communicate in Finnish. During the data collection, the participants were hospitalized for example due to labor induction. Polar H10 was used to collect abdominal ECG (Gilgen-Ammann *et al* 2019, Skála *et al* 2022), ECG sampling rate 130 Hz), and a hospital patient monitor (Philips Avalon FM30, Eenkhoorn *et al* (2024), ECG sampling rate 125 Hz) was used as the medical reference. The data collection lasted approximately 15 min per participant not including the instrumentation of the subject and the release process.

In this study, the Philips instrument was used to measure pregnant women's ECG ($mECG_{\text{ref}}$) and (fHR_{ref}). The $mECG_{\text{ref}}$ signal was measured with two ECG electrodes attached with the ECG adapter cable to the TOCO+ transducer, which were placed according to the manufacturer's instructions (one directly below the clavicle and near the right shoulder, the other in the left lower abdomen). The reference maternal heart rate (mHR_{ref}) was computed from the interbeat intervals (IBIs) detected from $mECG_{\text{ref}}$. The fHR_{ref} was measured with the Philips ultrasound transducer placed on the pregnant woman's abdomen. During the data collection, pregnant women were instructed to be still and in a half-sitting position, if possible.

Each data collection included three Phases of two-minute recording. Background information was also collected before starting the protocol. Those included information about duration of pregnancy, fetal position, and further observation as shown in table 2. The fetus's position was always head down which the researcher confirmed by palpation before placing the sensors.

Between the measurements, the placement of the H10 belt was changed to get the most comprehensive data possible (Phase 1: in the middle line of the body, Phases 2 and 3: approximately 20° to the left/right of the middle line of the body. Placement of other sensors was not changed. Figure 2 shows the

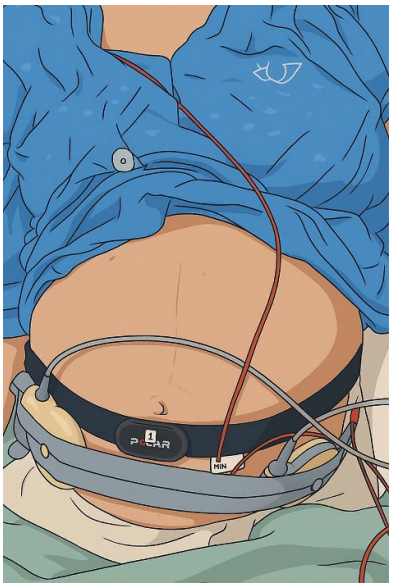

Figure 3. Polar H10 placement together with Philips FM30 TOCO+ sensors.

timing and events during a single session, which lasted around 60 min, and figure 3 shows the sensor placements.

Altogether, 19 pregnant women were recruited, of whom 12 were willing to participate in the study. Most common reason for refusing the participate in the study were frequent contractions associated with the first phase of labor, or the desire to focus on the approaching labor without unnecessary distraction. The first participant (ID 01) was considered as a pilot measurement and was excluded from the study due to inadequate placement of the H10 sensor.

3. Data pre-processing

All ECG signals were resampled at 200 Hz for the sake of consistency with the other databases and synthetic signals. The *aECG* is processed on a running window of 5 s with 50% overlap. We used MATLAB for the implementation of our codes (The MathWorks, Inc. MATLAB, Version R2024b, Natick, Massachusetts, USA, 2024).

Step 1
A 4th order Butterworth filter with cut-off frequencies between 0.5 Hz and 40 Hz was used on all ECG signals to attenuate the slow baseline wandering, breathing-induced amplitude modulations, and reduce high-frequency noise and 50 Hz power line interferences.

Step 2
The processed signals were normalized in amplitude using a running z-score window of five seconds with an overlap of 85%. The ECG signal resulting from Steps 1 and 2 was then used for further processing described in section 4.

Step 3
A signal quality index (*aSQI*) was then computed on *aECG*. The *aSQI* ranges from 0% to 100% and helps to select the signal with a reliable ECG waveform automatically. The *aSQI* can be viewed as a normalized SNR $aSQI_{\mathrm{SNR}}$ (dB) but however has different features as listed below. The later was estimated using the Savitzky–Golay filtering technique (Chakraborty and Das 2012) and was developed on the Physionet databases as follows:

(i) We computed the skewness (*SKEW*) and kurtosis (*KURT*) of *aECG*,
(ii) SQI_{SKEW} and SQI_{KURT} were computed from the 25% and 75% quartiles of *SKEW* and *KURT* with a piecewise linear approximation between 0% and 100%,
(iii) Finally, we define: $aSQI = \sqrt{SQI_{\mathrm{SKEW}}\ SQI_{\mathrm{KURT}}}$.

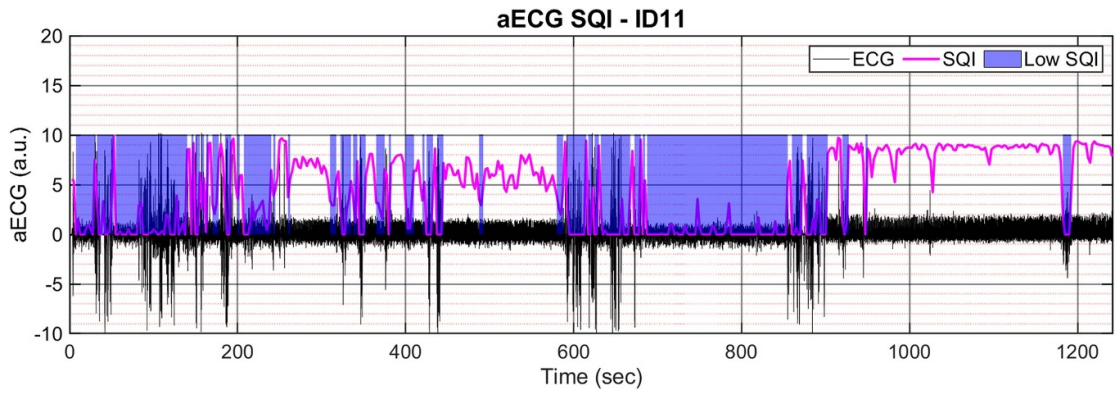

Figure 4. *aSQI* for subject 11 from the clinical data.

A threshold $T_{aSQI} = 40\%$ on *aSQI* enabled the selection of the most reliable signal's segments as shown in figure 4. The threshold was adjusted through visual inspection of PhysioNet and clinical data to optimize the number of reliable analysis segments. Although signal quality may sometimes be low, the *fQRS* remains visible and can still be utilized by our algorithm.

Figure 4 shows the *aSQI* during the three phases of the clinical data collection for relatively high to low average *aSQI*. The *aSQI* clearly detects short periods of low quality during the movement of the H10 belt between the protocol Phases. Phase 2 had the lowest *aSQI* while Phase 3 had the highest. During Phase 3, the *fQRS* complexes manifested clearly. The lowest *aSQI* was reported in the pre-Phase 1 during the electrode set up on the person.

4. *mHR* and *fHR* estimations

4.1. Branch 1: *mQRS* masking algorithm

Step 1

In Branch 1 (see figure 1) a modified PT algorithm (Pan and Tompkins 1985) was used to detect maternal m*R*-waves time locations $mR_{\text{ref}}(k)$ and $mR_{\text{PT}}(k)$ with k indexing each m*R*-wave[6]. The Philips instrument measures the *mECG* and therefore the $mR_{\text{ref}}(k)$ data correspond to the pregnant woman's heart. These data can then be used for the estimation of the raw pregnant woman reference IBI $mRR_{\text{ref}}(k) = mR_{\text{ref}}(k) - mR_{\text{ref}}(k-1)$.

From the location of the m*R*-waves, we then compute the width of each m*R*-wave by detecting the beginning and end of the wave from a drop of 50% of the wave maximum amplitude. Figure 5(a) shows an archetypal QRS complex together with ΔR and ΔQRS. This gives us a first estimation of the $mR_{\text{PT}}(k)$ width $\Delta mR_{\text{PT}}(k)$. We then filter out the $\Delta mR_{\text{PT}}(k)$ to ensure we are detecting the m*R*-waves of the pregnant woman and not the fetus. This step is necessary as the fetus's *R*-wave may have a similar amplitude to the pregnant woman's. We filter the $\Delta mR_{\text{PT}}(k)$ by imposing a certain range $\Delta mR_{\text{PT}}(k) \in [Min_{\Delta mR}, Max_{\Delta mR}]$. We chose $Min_{\Delta mR} = 40$ ms and $Max_{\Delta mR} = 100$ ms.

From $mR_{\text{PT}}(k)$, we compute the raw $mRR_{\text{PT}}(k)$. An IBI filter eventually detects and interpolate unphysiological $mRR_{\text{PT}}(k)$, resulting in the corrected series $mRR_{\text{PT}}(k)$. The IBI filter is based on a running median filter of length seven which detects IBI larger than 20% of the median of the seven previous IBIs, followed by an interpolation (Karlsson *et al* 2012). We use the median of the previous seven corrected IBI data for interpolation. This method is very efficient and allows for a fast real-time implementation without losing too much accuracy as compared to the more complex spline interpolation, for example (Benchekroun *et al* 2023). This technique, however, might not be accurate enough for heart rate variability analysis.

The $mHR_{\text{PT}}(k)$ is finally estimated using an running average window of $M = 10$ consecutive $mRR_{\text{PT}}(k)$ intervals. In a similar way, we estimated the mHR_{ref}.

Step 2

From the $\Delta mR_{\text{PT}}(k)$ data, we computed an hypothetical width for each QRS complexes $\Delta mQRS_{\text{PT}}(k) = 1.2\ \Delta mR_{\text{PT}}(k)$ and use it to construct a masking function $h(n)$ as described in equation (1). This

[6] The *R* wave is not the same as the *QRS* (i.e. *mQRS* and *fQRS*) introduced in section 2.1.

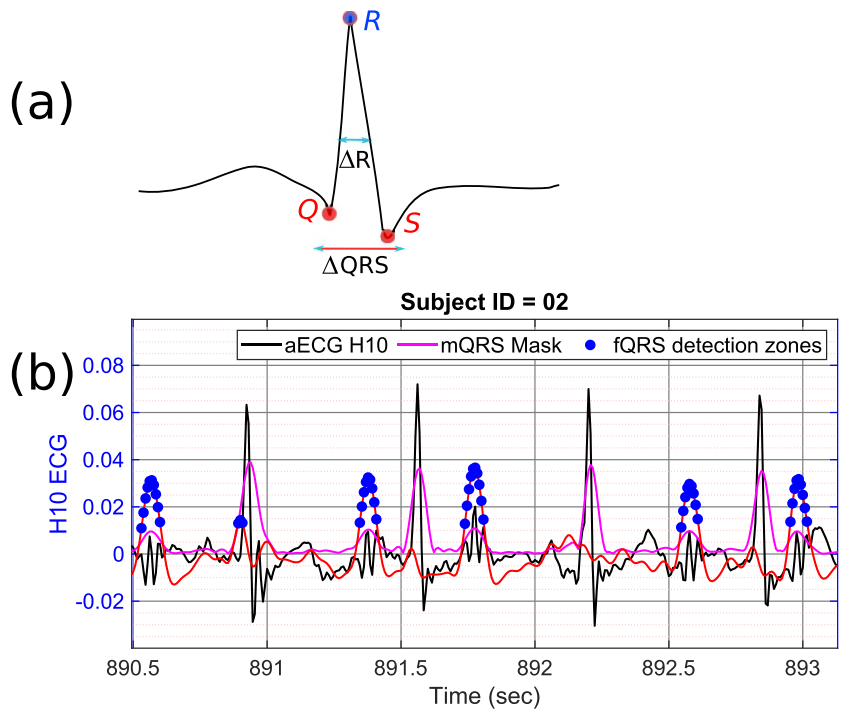

Figure 5. (a) Illustration of ΔR and ΔQRS for either fetus or maternal ECG. (b) Example of *mQRS* masking and detection of *fQRS* zones for Subject 02 in the clinical data.

procedure is similar to the one proposed by Jaros *et al* (2024), but we used a more straightforward approach based on a template masking function $h(n)$ with varying time support matching the $\Delta mQRS_{\mathrm{PT}}(k)$,

$$h(n) = \mathrm{e}^{-((n-k)/\Delta mQRS_{\mathrm{PT}}(k))^2}. \tag{1}$$

Each detected *mQRS* complex were then masked by $h(n)$ to produce another signal which should contain the *fQRS* complexes and some noise. The same procedure as for detecting the $mR_{\mathrm{PT}}(k)$ was used to: 1) detect the fetus $fR_{\mathrm{PT}}(k)$, 2) compute the $fRR_{\mathrm{PT}}(k)$, and 3) estimate the fHR_{PT}. We however used a different QRS width filter: $\Delta fR_{\mathrm{PT}}(k) \in \left[Min_{\Delta fR}, Max_{\Delta fR}\right]$. We chose $Min_{\Delta fR} = 20$ ms and $Max_{\Delta fR} = 60$ ms. Figure 5(b) illustrates the *mQRS* masking procedure on UTU subject 02 as well as the *fQRS* zones from which the maximum gives the f*R*-wave peak locations $fR_{\mathrm{PT}}(k)$.

A second IBI filter was especially designed for filtering the $fRR_{\mathrm{PT}}(k)$ and to best cope with the possible missed detection due to overlaps between the *mQRS* and *fQRS* (Matonia *et al* 2006). Figure 5(b) shows one example of overlap that will create an artificially long $fRR_{\mathrm{PT}}(k)$. An estimation of the fetus heart rate fHR_{PT} is then computed as the median of ten consecutive $fRR_{\mathrm{PT}}(k)$ which amounts to about 5 s at a heart rate of 120 bpm.

Step 3

A quality index $mfSQI_{\mathrm{PT}}$ is computed as the ratio of the maximum absolute amplitude of the $fQRS_{\mathrm{PT}}$ and the maximum absolute amplitude of $mQRS_{\mathrm{PT}}$[7]. When the mother-fetus QRS overlap last for multiple beats, the $mfSQI_{\mathrm{PT}}$ drops, and the corresponding data segment is flagged as low quality.

The $mQRS_{\mathrm{PT}}$ and $fQRS_{\mathrm{PT}}$, together with the signal quality $mfSQI_{\mathrm{PT}}$ are the main output of the Branch 1 which are the inputs for the Branch 3. The fHR_{PT} and fHR_{PT} are temporary estimations of the fetus and maternal heart rate which could be used when the signal qualities are high enough.

4.2. Branch 2: machine learning algorithm

Branch 2 (see figure 1) uses a machine learning approach based on a boosting decision tree classifier (Freund and Schapire 1996). The algorithm grows W trees with S splits sequentially using an iterative subsampling technique similar to the Random UnderSampling boosting algorithm (Seiffert *et al* 2008)

[7] Usually the maximum of $fQRS_{\mathrm{PT}}$ corresponds to the fR_{PT}.

Table 3. Features.

Feature	Description
$F^1_{(m,f,n)} = STD_{(m,f,n)}$	standard deviation of *mQRS*, *fQRS* and *Noise*
$F^2_{(m,f,n)} = KURT_{(m,f,n)}$	kurtosis of *mQRS*, *fQRS* and *Noise*
$F^3_{(m,f,n)} = SKEW_{(m,f,n)}$	skewness of *mQRS*, *fQRS* and *Noise*
$F^4_{(m,f)} = QRSWidth_{(m,f)}$	$\Delta mQRS_{\mathrm{PT}}$ and $\Delta fQRS_{\mathrm{PT}}$
$F^5_{(m,f)} = RPeak_{(m,f)}$	amplitude of the peak of *mQRS* and *fQRS*

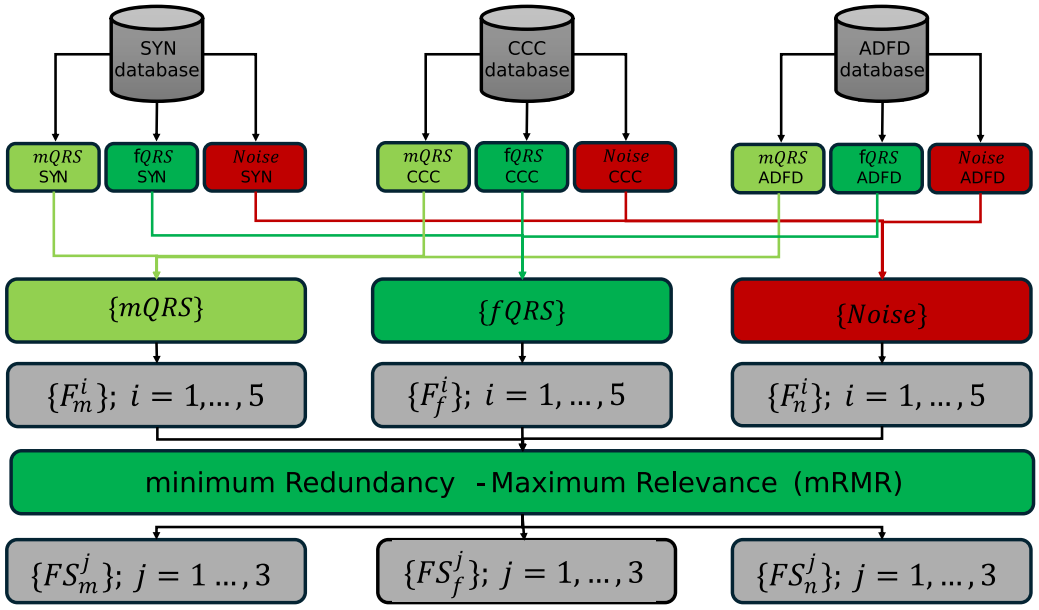

Figure 6. Flow chart of the feature extraction and selection.

whereby the performance of its performance is assessed at each added tree. The algorithm is described in more detail in appendix B.

Features' description: The classifier was trained on the Physionet and SYN data to recognize $M = 3$ classes: Class 1: *mQRS*, Class 2: *fQRS*, and Class 3: *Noise*. For each *mQRS*, *fQRS* and *Noise*, the features for the three classes $F^i_{(m,f,n)}$, for $i = 1, \ldots, 5$, in table 3 were computed and stored for further training, testing, and further validation on the clinical data (Huang *et al* 2024).

The noise features $F^{4,5}_{(n)}$ were computed on segments not containing *mQRS* and *fQRS*, and of duration $\Delta mQRS_{\mathrm{PT}}$.

Feature sets balancing: We collected a total number of feature samples $N = 76601$ from which Class 1 contains $N1 = 3970$ samples, Class 2 contains $N2 = 7980$ samples, and Class 3 contains $N3 = 64651$ samples. As the data are heavily unbalanced due to the Class 3 (QRS complexes can take between 7% to 20% of data points in a single heartbeat interval, depending on the heart rate), we randomly (uniform distribution) extracted Nx subsample from the data set with $N23 = p_2\ N2 + (1 - p_2)N3]$ samples from Class 2 and 3 with $p_2 = (N1 - N3)/(N2 - N3) = 1.0708$. The total number of samples for training and testing is $Nx = \lfloor N1 + N23 \rfloor = 7937$.

Feature selection: Feature selection was conducted using the minimum Redundancy—Maximum Relevance (Zhao *et al* 2019) criteria. The best three were retained for training the classifier: $SKEW_{(m,f,n)}$, $QRSWidth_{(m,f,n)}$, and $RPeak_{(m,f,n)}$ respectively by decreasing order of prediction power. By selecting three features per heartbeat interval, we compressed the information by a factor between 200/3 and 50/3 (for 200 Hz ECG sampling rate and heart rate between 60 bpm and 240 bpm). Figure 6 represents the feature extraction and selection flow chart. All *mQRS* patterns for all databases are denoted $\{mQRS\}$, and similarly for *fQRS* and *Noise*. The set of all input features is denoted $\{F^i_{(x)}\}$ with $x \in \{m, f, n\}$ for $i = 1, \ldots, 5$. The selected features are denoted $\{FS^j_{(x)}\}$ for $j = 1, \ldots, 3$.

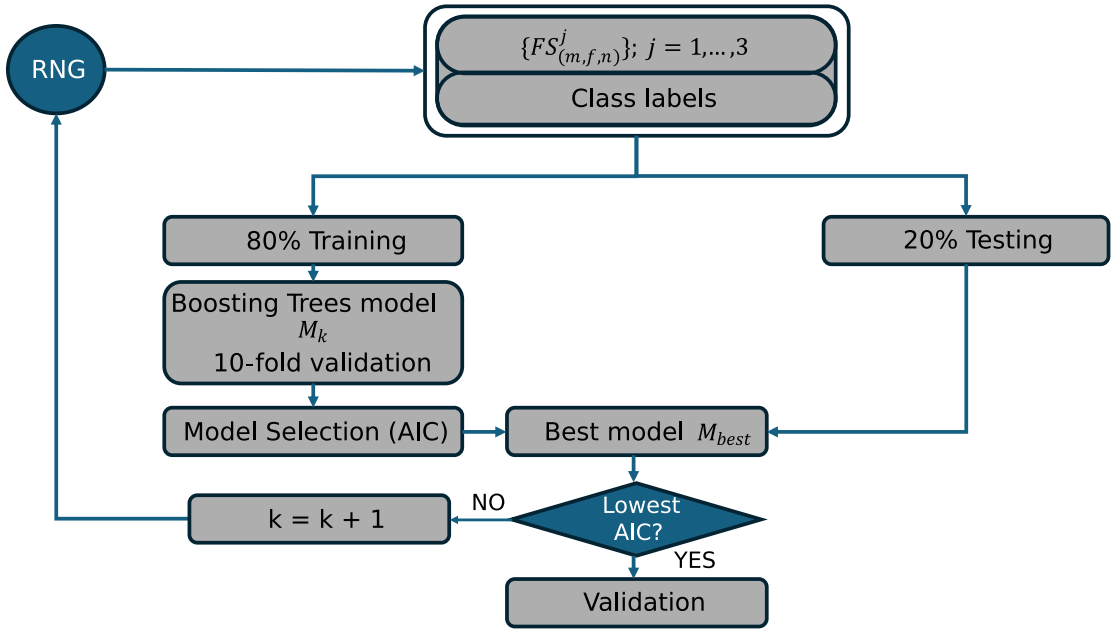

Figure 7. Flow chart of the machine learning with model selection.

Table 4. Machine Learning Algorithm Statistics at optimal points.

Metrics (%)	$mQRS_C$	$fQRS_C$	*Noise*
$Sensitivity^{Optim}$	96.2	87.0	96.4
$Specificity^{Optim}$	94.3	97.5	97.3
$Precision^{Optim}$	89.4	95.0	94.7

Classifier: The true class labels for training set of *fQRS* came from the annotations in the CCC and ADFD databases, while those for the training set of *mQRS* came from the Branch 1 as there are no annotations for the pregnant woman QRS locations. When using the SYN data generator, the *a priori* known *mQRS* and *fQRS* were used as true labels. A 10-fold cross-validation was conducted on a dataset randomly divided into 80% training and 20% testing feature subsets. The validation was done on four best subjects' recording in the clinical database.

The chosen machine learning model is based on a boosting architecture with an increasing number of decision tree classifiers. As this work intends to embed the code into a low-power module, we optimized the model parameters: number of learners W and splits S by a grid search using a similar concept as the Akaike criteria $AIC_{(m,f,n)}$ (Akaike 1998, Cavanaugh and Neath 2019):

$$AIC_{(m,f,n)}(W,S) = -Sensitivity_{(m,f,n)} + \lambda_{(m,f,n)}\ Complexity/N1 \tag{2}$$

with m, f, and n refer to mother, fetus and noise. The complexity was defined as: $Complexity = W\,S$[8]. A penalty term $\lambda_{(m,f,n)}$ for the complexity was introduced and was adapted to each Class: $\lambda_m = 0.3$, $\lambda_f = 0.6$, $\lambda_n = 0.3$. This specific Akaike model selection (AIC), whereby the Sensitivity and Complexity increase with W and S, is derived under particular assumptions in appendix A.

Figure 7 shows the machine learning training, testing and model selection architecture. The 10-fold cross-validation boosting tree model at iteration k is denoted $\mathcal{M}_k$ while scanning all possible values of S and W using $AIC_{(m,f,n)}(W,S)$. The lowest value of $AIC_{(m,f,n)}(W,S)$ is retained and the best model $\mathcal{M}_{\text{best}}$ is retained for validation.

Table 4 summarizes the optimal performances corresponding to the minimum of the AIC curves as found in appendix A. The *fQRS* detection $Sensitivity^{Optim}$ and the *mQRS* detection $Precision^{Optim}$ have the lowest value. The classifier detects the $mQRS_C$ and $fQRS_C$ complexes from which we locate the peak *R*-waves mR_C and fR_C, compute the corresponding IBI and the heart rate fHR_C as described in section 4.1. The $mQRS_C$ and $fQRS_C$ together with the signal quality $mfSQI_C$ are the main output of the Branch 2

[8] The Complexity does not take into account the time, energy and resources required for training the algorithm.

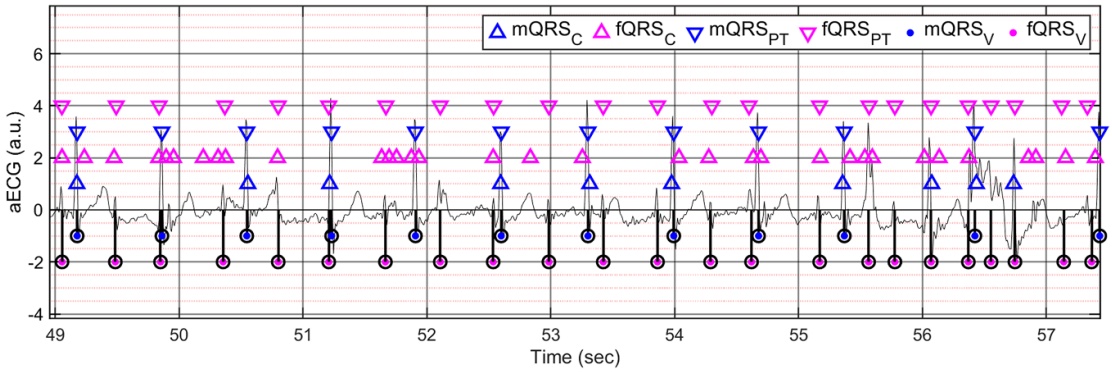

Figure 8. Example of $mQRS_{PT}$ and $fQRS_{PT}$ (Branch 1), $mQRS_C$ and $fQRS_C$ (Branch 2), and $mQRS_V$ and $fQRS_V$ (Branch 3) detections for Subject 05 in the clinical data.

which are the inputs for the Branch 3. As in Branch 2, the fHR_C and fHR_C are temporary estimations of the fetus and maternal heart rate which could be used when the signal qualities are high enough.

4.3. Branch 3: dynamic fusion

As mentioned in sections 4.1 and 4.2, fHR_{PT} and fHR_C can be used to estimate a final fHR using data fusion. The fusion would be based on the following weighted average formula:

$$fHR = \frac{mfSQI_{PT}\ fHR_{PT} + mfSQI_C\ fHR_C}{mfSQI_{PT} + mfSQI_C}. \tag{3}$$

The approach based on equation (3) works well but lack robustness when the $mfSQI_{PT}$ and/or $mfSQI_C$ are low. The solution that we have adopted uses the Viterbi algorithm (Viterbi 1967) as shown in figure 1. The Viterbi algorithm uses the $fQRS$ and $mfSQI$ from the Branch 1 and 2 to estimate an fHR. The dynamic aspect is automatically implemented by the Viterbi decoding which uses past data to predict the $fQRS_V$ and finally estimate the final fHR using the same IBI filtering and HR estimation procedure as before.

As we recall, our implementation is based on detecting a candidate $fQRS(k)$ at time $fR(k)$. The quality indices $mfSQI(k)$ from Branch 1 and 2 are interpreted as probabilities that the Branch 1 or 2 correctly identify a fetus QRS at time $t_k = k/F_s$. Transitions are allowed from a previous candidate i at time t_i to a later candidate j at time t_j with $fRR(i,j) = fR(j) - fR(i) \triangleq fRR(i)$ only if $fRR(i) \in [Min_{\Delta fR}, Max_{\Delta fR}]$. Indexing candidates $fR(k)$ in chronological order by $k = 1,\ldots,N$, N being the total number of beats detected in both Branches, we select a subset of these candidates, that maximize the Viterbi score (the details of the Viterbi are described in C).

Figure 8 illustrates the detection of $mQRS_C$ and $fQRS_C$ *zones* from Branch 2, along with $mQRS_{PT}$ and $fQRS_{PT}$ from Branch 1 and finally the Viterbi consensus decoding $mQRS_V$ and $fQRS_V$ from Branch 3.

5. Results

5.1. Physionet databases

The Physionet ECG data contains four channels and 5 + 75 subjects, totaling 320 ECG recordings. We processed each channel separately. For each channel, we computed the following metrics: 1) the Mean Absolute Percentage Error ($MAPE$) between the computed fHR and the reference annotations fHR_{ref}, 2) $aSQI_{SNR}$, and 3) $mfSQI_{SNR}$. Figure 9 shows the different metrics using heat maps for each recordings, with the last five $[76,77,78,79,80]$ being from the ADFD database.

The white dots in figure 9(a) indicate when $MAPE \leqslant 5\%$ ($MAPE_5$) and the red dots when $MAPE \leqslant 10\%$ ($MAPE_{10}$). 24% of the 320 recordings (all channels included) belong to $MAPE_5$ and 35% $MAPE_{10}$. Out of the 80 subjects (at least one channel/subject), 44% belong to $MAPE_5$ and 76% belong to $MAPE_{10}$.

Figure 9(a) show small $MAPEs$, as indicated by the high density of white dots in figure 9(a). This corresponds to higher $mfSQI_{SNR}$ in figure 9(a), which appears to be a sufficient condition for achieving a low $MAPE$. In addition, figures 9(b) and (c) suggest an apparent inverse relationship between $aSQI_{SNR}$ and $mfSQI_{SNR}$, as can be seen in figure 10(f), except for the ADFD set. The reason is that the $mfSQI_{SNR}$ is a relative quality measure of the maternal and fetal QRS amplitudes and $aSQI_{SNR}$ intrinsically contains

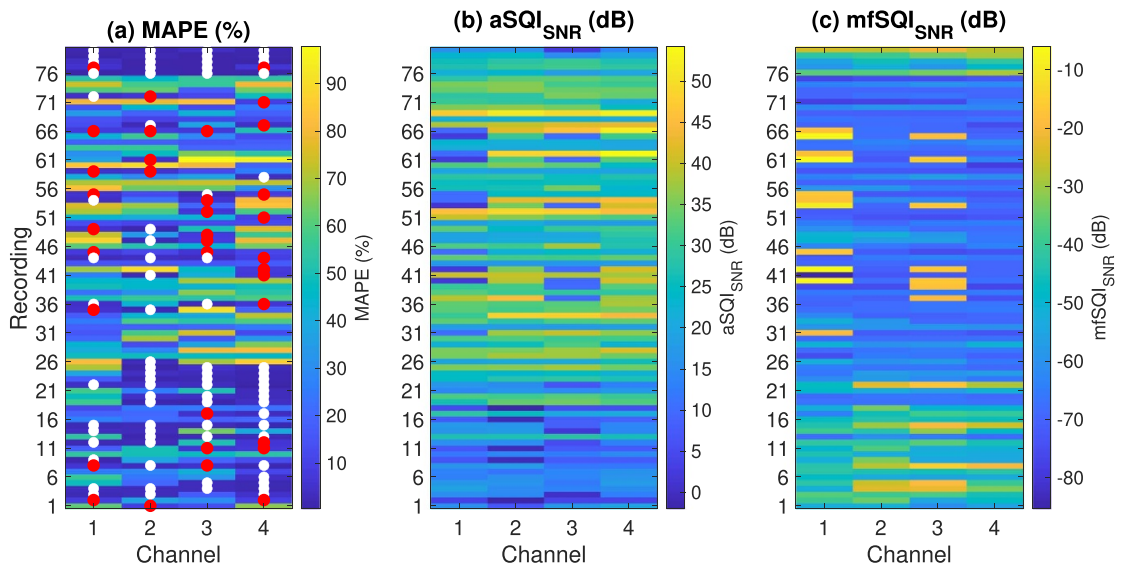

Figure 9. CCC and ADFD statistics: (a) *MAPE*, (b) $aSQI_{SNR}$, and (c) $mfSQI_{SNR}$.

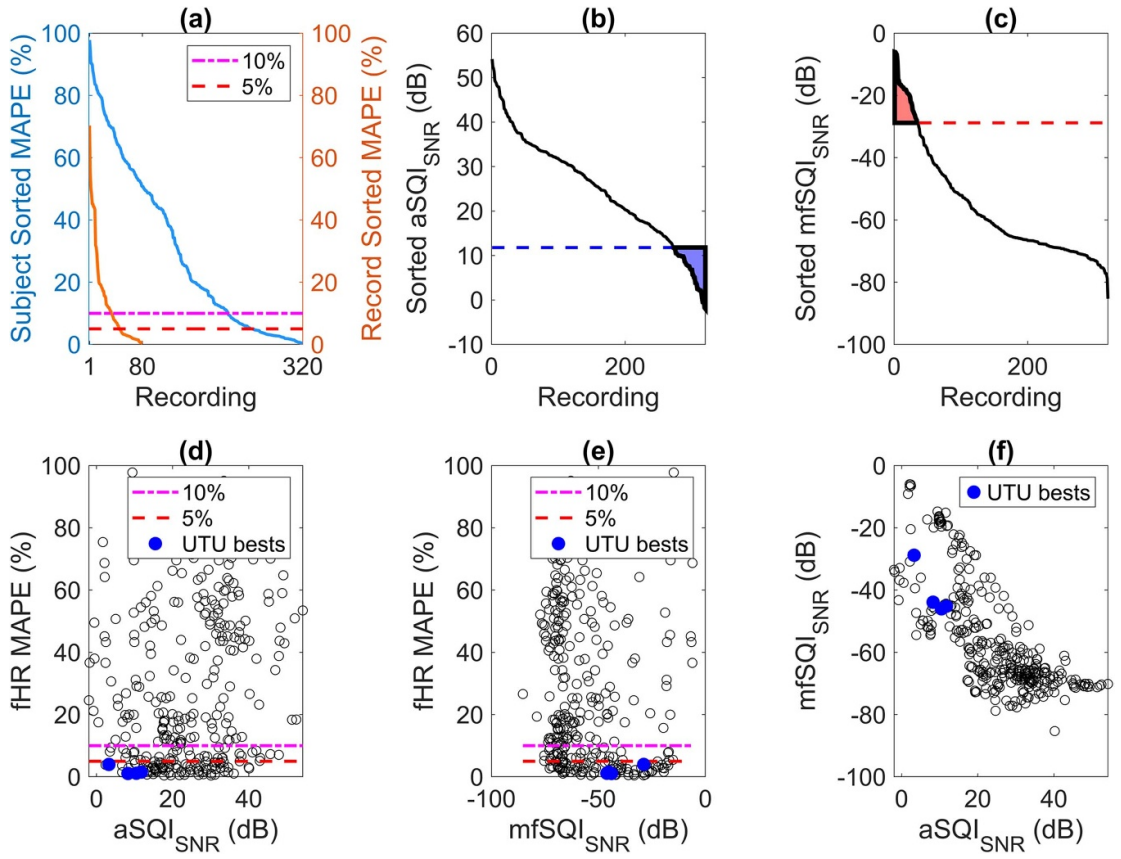

Figure 10. Physionet statistics: the top row contains the sorted metrics, and the bottom row contains the relationships between them.

the *fECG* and noise contributions *Noise*. Figure 10(d) demonstrates that there is no relationship between *fHR MAPE* and $aSQI_{SNR}$.

Figure 10(a) further confirms the difference between single-channel and multi-channel performances as noticed above from the percentage of MAP_xE. Figure 10(e) shows a sharp rise of *fHR MAPE* when $mfSQI_{SNR} \lessapprox -70$ dB which is a lower threshold below which almost no *fQRS* can be detected with our algorithm.

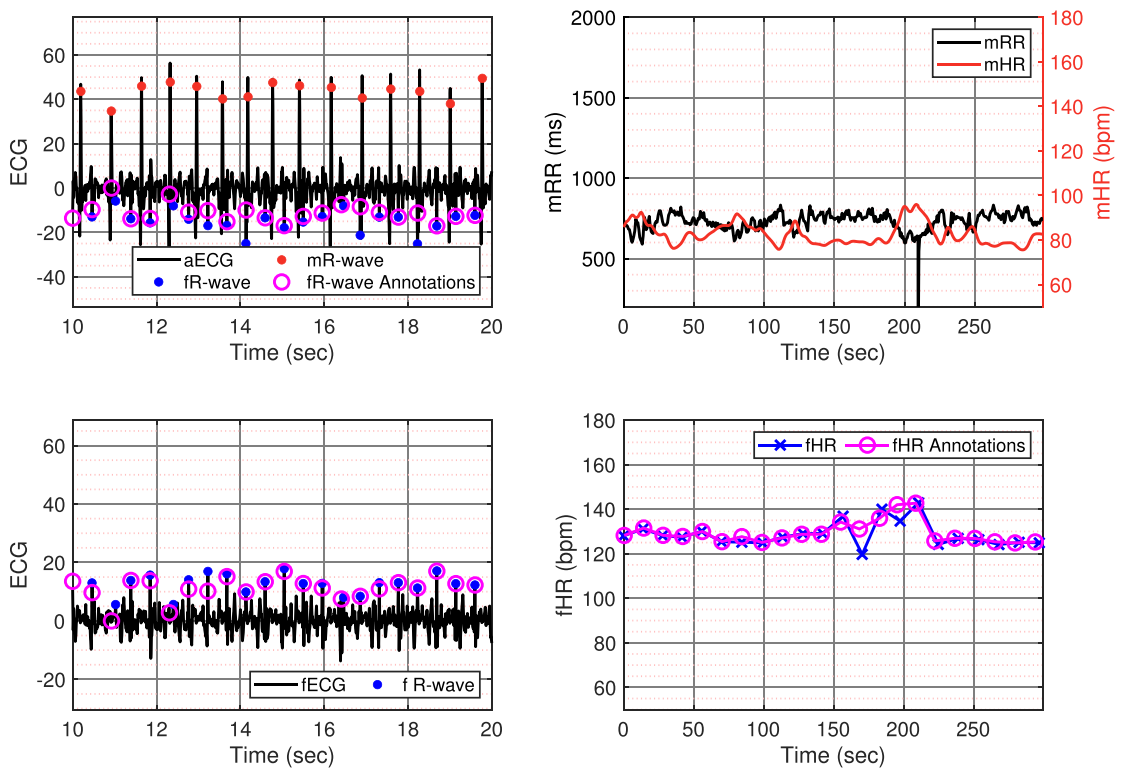

Figure 11. ADFD subject r01, channel 1, analysis result.

Table 5. Machine learning algorithm statistics at optimal points for the ADFD and CCC database. Average (*AV*) and standard deviation (*SD*) across all records are reported.

	CCC		ADFD	
Metrics (%)	*AV*	*SD*	*AV*	*SD*
*Sensitivity**	93.9	3.6	97.4	0.4
*Specificity**	93.8	3.2	99.8	0.1
*Precision**	82.2	8.7	97.2	1.9

A detailed analysis on subject r01, channel 1, from the ADFD database is shown in figure 11. In this recording we have $aSQI_{\mathrm{SNR}} = 22$ dB, $mfSQI_{\mathrm{SNR}} = -36$ dB and $MAPE = 0.45\%$. This example shows the robustness of our algorithm even when the *fECG* is very noisy. The *fHR* shows a slow rise until about 200 s followed by a sharp fall around 220 sec, well captured by the algorithm. Interestingly, the *mHR* displays a sharp rise just prior to 200 s falling back to a sort of baseline around 220 section During this event, the *mHR* and *fHR* are positively correlated. In this very good quality signal,m we can thus expect to perform detailed analysis of rise and fall of the *fHR* and eventually the coupling between *mHR* and *fHR*.

Machine learning performance metrics includes: sensitivity, specificity and precision. These metrics were computed for the *fQRS* detection and are reported in table 5. Due to the multi-channel nature of these databases, we have summarized the statistics as the average of the best channel for each record. The best channel n for the record i was selected as the channel $n(i) = n^*(i)$ which maximizes the Youden index (Fluss *et al* 2005) $Y(i,n) = Sensitivity(i,n) + Specificity(i,n) - 1$:

$$n^*(i) = \arg\max_n Y(i,n). \tag{4}$$

We then define the best $Sensitivity^*(i) = Sensitivity(i, n^*(i))$, $Specificity^*(i) = Specificity(i, n^*(i))$ and $Precision^*(i) = Precision(i, n^*(i))$ for each record i. Their average and standard deviations for all records are displayed in table 5 where the ADFD and CCC database statistics are shown separately as they have different quality as shown in figures 7(b) and (c).

We collected the *Sensitivity*, *Specificity* and *Precision* of all 80 recordings for each channels sorted by decreasing value of the Youden index from the CCC and ADFD databases. Figure 12 shows the distribution of these three statistics. The statistics reported in table 5 are equivalently computed from $ChSort_1^*$.

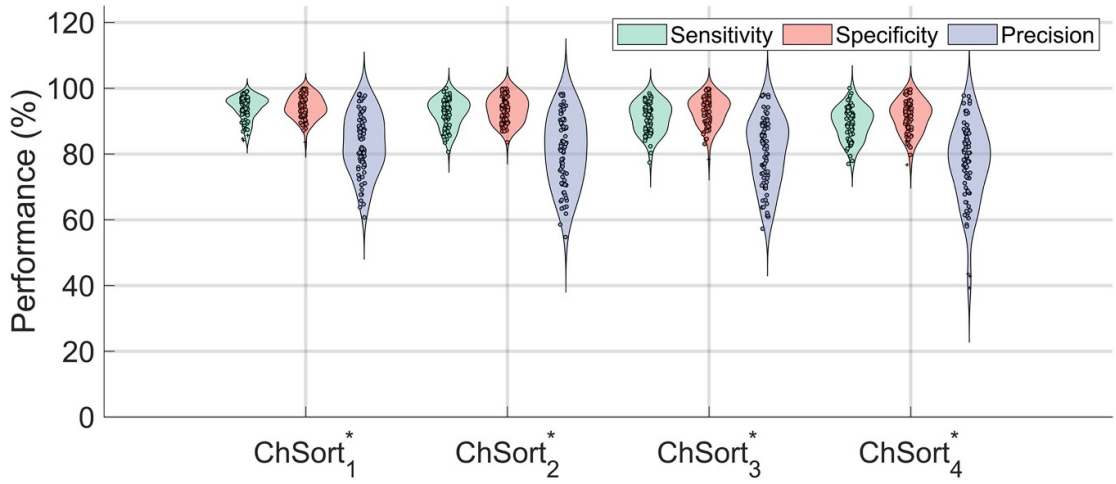

Figure 12. CCC and ADFD *Sensitivity*, *Specificity* and *Precision* distributions sorted by decreasing Youden index for each channel.

Table 6. Average statistics per subject. *MAPE* and *Pa* in percent and SQI_{SNR} in decibels.

ID	*mHR MAPE*	*fHR MAPE*	*Pa*	$aSQI_{\mathrm{SNR}}$	$mfSQI_{\mathrm{SNR}}$
2	2.7	4.0	87.0	3.3	−28.8
5	3.6	1.1	77.0	8.3	−43.9
9	2.5	1.1	80.0	10.5	−46.0
11	4.1	1.5	74.4	11.8	−45.2

Figure 12 shows that the average and standard deviations of *Sensitivity* and *Specificity* and *Precision* are decreasing from $ChSort_1^*$ to $ChSort_4^*$.

Wilcoxon rank sum tests across channels for *Sensitivity* indicates that all pairwise channels achieve statistical significance with $p \leqslant 0.033$. Similar tests on *Specificity* show significance between $ChSort_{1,2,3}^*$ and $ChSort_4^*$ with $p = 0.0003$, $p = 0.0056$ and $p = 0.021$ respectively. The *Precision* reached only significance between $ChSort_1^*$ and $ChSort_4^*$ at $p = 0.0014$. These results gives us some insights into the usefulness of multichannel ECG systems for *fHR* estimation and are discussed in section 6.

5.2. Clinical data

Of the 11 participants' recordings, seven were excluded for one or more of the following reasons (see table 2): 1) failure in transferring data to the recording device, 2) malfunctioning of certain instruments, 3) excessive movement by some expectant mothers resulting in substantial artifacts, and 4) difficulty in properly adjusting the H10 sensor for one participant with a small abdominal circumference.

The four best subjects (2,5,9,11) were analyzed in detail. The *mHR* data were resampled to 1 Hz and synchronized with the reference. Additionally, the *fHR* data were also computed as an average over 30 s segments. We collected 80 segments for all subjects after discarding those where $mfSQI \leqslant Tf$ with $Tf = 40\%$ and $aSQI \leqslant Ta$ with $Ta = 70\%$. This selection resulted to a total of about 45 min of reliable data which amount to $45/(4 \times 15) = 0.75$ (15 min recording time per subjects for four subjects) percent. The percentage of reliable segments is called *availability* (*Pa*). The maternal and *fHR MAPE* were computed on reliable segments. Table 6 summarizes the accuracy statistics together with the average $aSQI_{\mathrm{SNR}}$ and $mfSQI_{\mathrm{SNR}}$.

We notice that the *mHR MAPE* is significantly higher than *fHR MAPE*. The reason might be due to the sometime low quality of the Philips ECG due to ECG electrode movement, tension in the electrode wires or electromagnetic interferences from the recording machines.

Figure 13 shows the scatter and Bland-Altman plots of the *fHR* and fHR_{Ref} on all four good subjects and all segments. The lower and higher 95% Limit of Agreement (± 1.96 *SD*) are $[-15, 19.5]$ bpm with an average bias $\langle \Delta fHR \rangle = 1.97$ bpm where $\Delta fHR = fHR - fHR_{\mathrm{Ref}}$. The lower and higher percentage of data outside the standard deviation of ΔfHR are $[5.0, 12.5]\%$. The coverage probability when expressed as a percentage CP_ϵ (Barnhart *et al* 2007) (i.e. the probability of measurement error falling within a predefined bound ϵ bpm) were $CP_5 = 76.2\%$ and $CP_{10} = 87.5\%$.

Table 6 shows that the maximum $aSQI_{\mathrm{SNR}} = 11.8$ dB, which corresponds to about 15% of the worst data in the Physionet databases (see figure 8(b) blue region), and the maximum $mfSQI_{\mathrm{SNR}} \approx -28.8$ dB,

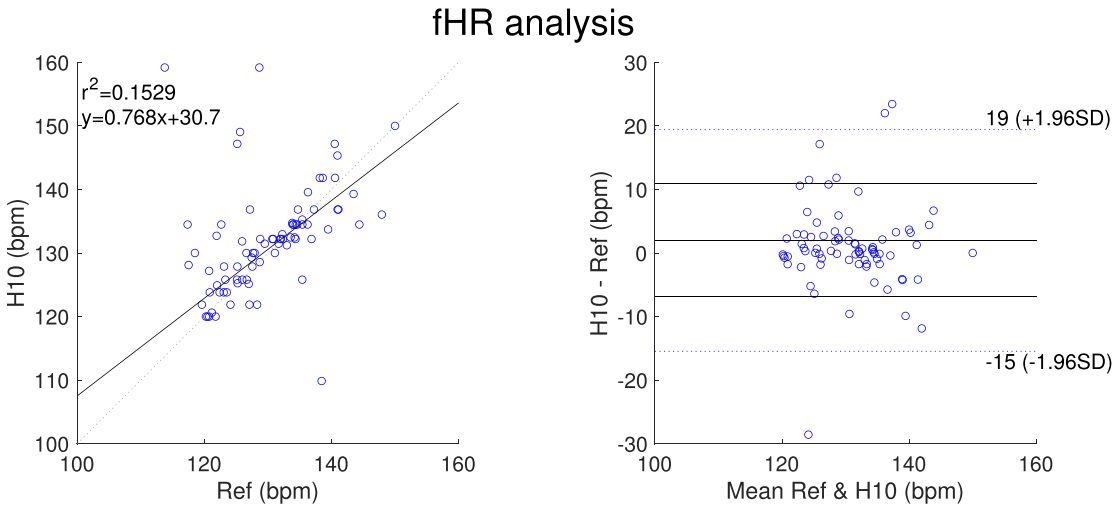

Figure 13. Bland Altman analysis of *fHR* and fHR_{Ref} on all four good subjects and all segments.

which corresponds to about 10% of the best data in the Physionet databases (see figure 8(c) red region). This suggest that dry-electrodes could potentially be used for *fECG* analysis upon further denoising.

6. Discussion

For the sake of clarity, the discussion has been organized into distinct sections addressing the technological and clinical aspects of this study.

Algorithm: This pilot study highlights the potential of a hybrid algorithmic architecture, in which the integration of machine learning and data fusion enhances the reliability of *fHR* estimation. The architecture is designed for portability to low-power embedded systems, thereby supporting secure data management. The machine learning component employs decision tree boosting, a method well recognized for its adaptability across diverse classification tasks. To optimize performance, we developed a novel model selection procedure tailored specifically for machine learning applications. This procedure, based on the Akaike information criterion, is broadly applicable to other machine learning problems. Importantly, the selection of optimal model parameters enables adjustment of sensitivity and specificity, as the Akaike criterion exhibits relative flatness in the parameter space surrounding its minimum. The robustness of our approach is further demonstrated by the autonomy of Branches 1 and 2, which can independently generate *fHR* estimates in the event of failure. Moreover, Branch 3 provides a consensus estimate when partial reliability is achieved, guided by our signal quality indices.

Dry electrodes: Despite continuous progresses, the use of dry electrodes remains a challenge as they promote noise. This is especially true for those using fabric as they also face the issues of washability and long-term fabric-skin irritability and comfort (Searle and Kirkup 2000, Galli *et al* 2021, Joutsen *et al* 2024, Kumar *et al* 2025). The carbon-base electrodes used in Polar H10 might be a good alternative.

Noise: Noise and maternal fetal QRS overlaps are challenging situations for classic signal processing techniques. A recent work in progress on simulated single channel *aECG* data by Orvas *et al* try to address these issues by leveraging the power of Complex-valued U-Nets (Orvas *et al* 2025) on time-frequency decomposition of the *aECG*. This work also aims at demonstrating the possible development of dry-electrode platforms. They report an average *Sensitivity* = 72.1% on the CCC database compared to our *Sensitivity* = 93.9%.

Deep Learning: Machine or Deep learning algorithms are being deployed in the most recent *fHR* algorithm based on abdominal maternal ECG. Excellent results were obtained for example by Zhou *et al* (2024, 2025) on the ADFD and CCC databases where they achieved average *Precision* = 99.8(93.7)% respectively as compared to our best results *Precision* = 97.2(82.2)%. The complexity of DL algorithms

pushes their deployment on cloud computers with an increase in communication bandwidth which put them in an other strategic category altogether compared to our reduced complexity goal.

Portable solution: The a-priori fixed electrode set-up might limit its use in home settings as the geometry might not be optimal creating false alarms and stress for the expectant mother. An alternative solution would be to have a movable system which, when displaced on the belly, would automatically detect the best position by informing the mother by a sound when the best *mfECGSQI* has been found. This portable unit and its biofeedback would potentially result in lowering the false alarms, higher *fHR* accuracy and diminished mother's stress.

Multichannel aECG: The Physionet databases contains four channels. As our algorithm is based on a single channel approach, it is without doubts that a multichannel algorithm would eventually perform better. The above analysis points to the fact that there is generally one channel which shows significantly better *Precision* with less dispersion in the *Sensitivity* and *Specificity*.

Clinical data: We noticed that for some of our clinical subjects, the *fECG* was not visible despite moving the H10 sensor around the abdomen. This observation was confirmed by low *mfSQI*. The electrical activity of the fetus's heart propagates throughout the pregnant woman's uterine tissues and liquids such as the amniotic fluid and the placenta (Keenan *et al* 2018, Smith *et al* 2018), which should facilitate its propagation. However, it seems that the cause of the *fECG* disappearance, in addition to various noise interferes (Matonia *et al* 2020), is not explicitly linked to a lack of either of these fluid quantities or location. Instead, we believe that the cause of such behavior is mainly due to the relative orientation of the fetus's heart dipole with respect to the H10 electrodes (Rooijakkers *et al* 2014). Another potential cause would be the presence of excessive vernix caseosa on the fetus's body (Uv *et al* 2025). Future *aECG* systems must have the capacity to automatically detect the best electrode orientation to maximize the $mfSQI_{SNR}$. we conclude that, despite the medium H10 ECG quality due to the unusual sensor setting, the ECGs of the four best clinical subjects were of relatively good quality. We expect to improve the accuracy and reliability of the approach if a multi-electrode architecture would be deployed, still keeping its user friendliness.

Clinical analysis: Since one of the key challenges in home health monitoring is the reliability and accuracy of the measurements, the results of this study were also promising from a health care perspective. Although the accuracy of the used method is still limited based on this pilot study, the results form a basis for further development of both measuring devices and measurement technique.

Before such technology can be used as part of pregnancy care services, the significance of the results should also be considered from the perspective of clinical research. Since one of the advantages of remote monitoring is considered to be the possibility of continuous monitoring (Saarikko *et al* 2020), it should be noted that current treatment and monitoring practices are based on short-term measurements as part of health care visits (Rowe *et al* 2020). Finding an indiscrete, easy-to-use, and reliable device for *fHR* monitoring, for example, enables long-term data collection even between health care visits. Understanding normal fetal behavior during pregnancy and changes in heart rate are prerequisites for the use of continuous pregnancy monitoring.

From the perspective of future health care, this type of technology would enable closer monitoring of pregnancies without the need for hospital stays or multiple visits to health care (Atkinson *et al* 2023). For example, inducing labor is becoming increasingly common (Kruit *et al* 2022), and technology would enable hospital-level monitoring even at home. However, as pregnant woman and fetus are strongly connected, it is essential that remote monitoring system could combine *fHR* to other fetal and maternal parameters (Likitalo *et al* 2025). For example, when assessing possible fetal hypoxic stress during labor, in addition to the heart rate, the relationship between possible decelerations in *fHR* and uterine contractions should also be taken into account (Ross 2011).

Since both the health care system as an implementation context and pregnancy as a health condition are extremely complex, it is therefore advisable for remote monitoring developers to closely collaborate with future end-users to identify the existing clinical needs (Zander *et al* 2023, de Plaza *et al* 2023). For example, in terms of acceptability, it is already known that pregnant women consider that devices used for remote monitoring should be as small and easy to use as possible (Stricker *et al* 2025). The device used in the study would be well suited to these needs, which supports its further development in the future.

7. Conclusion

This study reports the proof of concept that the Polar H10 sensor can capture the fetal ECG signal in some conditions. An algorithm based on a hybrid use of machine learning and classic digital signal processing could delineate the fetus's QRS from the pregnant woman's ECG signal and estimate the *fHR*. For the four best subjects, the performance is comparable to the best results from the Physionet data.

Our approach can be integrated into low-power devices without relying on complex deep learning algorithms, which helps avoid consuming large computing and energy resources for training. Researchers may also prefer our hybrid approach because only a limited number of abdominal maternal ECGs and reference fetal ECGs exist for training deep learning algorithms. Although synthetic ECG models can generate a considerable amount of data, they can never reproduce all the diverse aspects of real-life signals.

8. Study limitations

The Clinical data consist of 11 subjects, out of which seven were unusable for various technical reasons. The gestational age was limited between 37 and 41 weeks which could potentially restrict the use of such *aECG* technology. Further tests should be performed at an earlier gestational time. Therefore the statistical results on the clinical data must be taken with care. The Polar H10 sensor used in this study was not optimized for *fHR* estimation, except for the enlargement of the sensor belt. Future dry electrode *aECG* systems must improve the belt and ECG patch sensor placement and number.

9. Ethical statement and consent

Our research was conducted in accordance with the principles embodied in the Declaration of Helsinki and in accordance with local statutory requirements. The ethical approval for the study was received from the Ethics Committee of the Wellbeing Services County of Southwest Finland in August 2024 (VARHA/12 935/13.02.02/2024). The study was also registered to clinical trials.gov before starting the data collection (NCT06683183).

The eligible pregnant women meeting the inclusion criteria were identified in co-operation with health care professionals working in the ward and they were provided with information about the study and the privacy statement. Pregnant women willing to participate in the study gave their written informed consent before starting the data collection.

Data availability statement

The data cannot be made publicly available upon publication because they contain sensitive personal information. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

Data collection was performed at the Maternity Hospital of the Wellbeing Services County of Southwest Finland. The European Project NewLife provided funding. This project is supported by the Chips Joint Undertaking (Grant Agreement No. 101095792) and its members Finland, Germany, Ireland, the Netherlands, Sweden, and Switzerland. This work includes top-up funding from the Swiss State Secretariat for Education, Research and Innovation (SERI). In addidion, Finnish partner receives funding from Business Finland (4136/31/2022).

Susanna Likitalo performed the recording and participated in the writing of the paper. Arman Anzanpour installed the Philips instruments and guided the nursing team for proper recording, configured the web services, and data transfer protocols. Anna Axelin managed the University of Turku team and supervised the project's operations. Tommi Jaako prepared the Polar devices, instructed the University of Turku team on how to use them, and managed the data transfer from the hospital to Polar cloud systems for further analysis. Patrick Celka developed, tested, and validated the algorithms and participated in the writing of the paper.

Author contributions

S Likitalo 0009-0009-9413-2090
Investigation (lead), Methodology (equal), Writing – original draft (equal)

A Anzanpour
Investigation (supporting), Software (supporting), Writing – review & editing (supporting)

A Axelin
Supervision (lead), Writing – review & editing (equal)

T Jaako
Investigation (supporting), Software (supporting)

P Celka
Data curation (lead), Formal analysis (lead), Software (equal), Writing – original draft (equal)

Appendix A

Model selection is crucial to avoid overfitting in regression and machine learning. The best model $\mathcal{M}_{opt}$ out of a series of models $\mathcal{M}_n$ for $n = 1, \ldots, C$ is the one that minimizes a cost function. Several cost functions can be used, and we chose Akaike's one. The derivation of the Akaike criteria (AIC) for machine learning is derived from statistical analysis of a multiclass Gaussian Bayes optimal Linear Discriminant Analysis Model. Murphy (2012). The AIC criteria is defined as Akaike (1998), Cavanaugh and Neath (2019):

$$AIC = -\log\left(\mathcal{L}_{\max}\left(\hat{\theta}|y\right)\right) + \beta\ K \tag{A.1}$$

where $\mathcal{L}_{\max}(\hat{\theta}|y)$ is the maximum Likelihood of the model $\mathcal{M}_{\hat{\theta}}$ given an output y and λ is a free parameter. In our classification problem, let us assume we have C classes and a feature vector $\mathbf{y} \in \mathbb{R}^N$ which is produced by a model $\mathcal{M}_n$. We define $\Delta_{i,j}$:

$$\Delta_{i,j} = 2\left(\log\left(\mathcal{L}\left(i,i\right)\right) - \log\left(\mathcal{L}\left(i,j\right)\right)\right) \tag{A.2}$$

with $\mathcal{L}(i,n) = \mathcal{L}(y_i|\theta = n)$ the likelihood of model $\mathcal{M}_n$ given the feature y_i. $\Delta_{i,j}$ is the Mahalanobis distance between the two distributions and is directly related to the Kullback–Leibler (KL) divergence between the two models (Cavanaugh and Neath 2019). We can thus estimate the average KL divergence:

$$\Delta_i = 2\mathbb{E}_j\left(\log\left(\mathcal{L}\left(i,i\right)\right) - \log\left(\mathcal{L}\left(i,j\right)\right)\right) \tag{A.3}$$

and the AIC for a model $\mathcal{M}_i$ is approximated by:

$$AIC_i = -\Delta_i + \beta\ K. \tag{A.4}$$

For a Linear Discriminant classifier under Bayes optimization, it has been shown (Murphy 2012) that the sensitivity for a model $\mathcal{M}_i$ is given by:

$$Sensitivity_i = \Phi\ (\Delta_i/2) \tag{A.5}$$

with $\Phi(x)$ the cumulative distribution function of x. Therefore, the AIC for the model $\mathcal{M}_i$ is:

$$AIC_i = -2\Phi^{-1}\left(Sensitivity_i\right) + \beta\ K. \tag{A.6}$$

Assuming that the Sensitivity will be in the interval $[0.5, 1[$, equation (A.6) can be linearly approximated by:

$$AIC_i \approx -\alpha\ Sensitivity_i + \beta\ K + \gamma. \tag{A.7}$$

For this work and without loss of generality, we will simplify the equation (A.7) by imposing $\alpha = 1$ and $\gamma = 0$ as it will not influence the minimization of the AIC.

From equation (2), figure A1 display a heat map of each $AIC_{(m,f,n)}(W,S)$ which clearly indicates the zones where the $AIC_{(m,f,n)}(W,S)$ are minimized. All three $AIC_{(m,f,n)}(W,S)$ have minimum values in the range of $W \in [2,40]$ and $S \in [2,10]$.

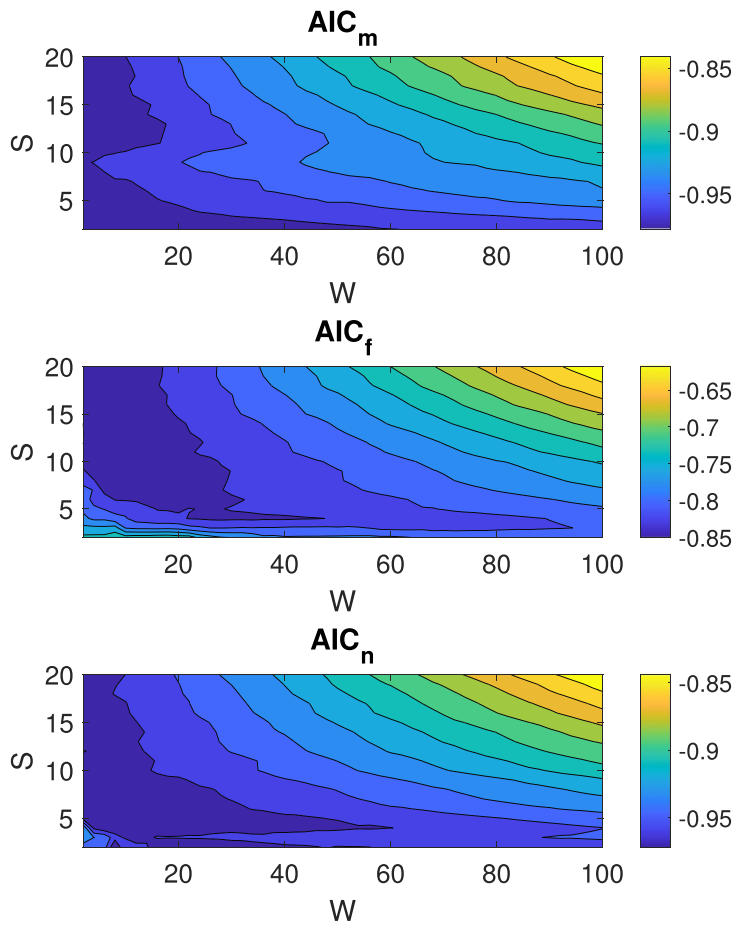

Figure A1. $AIC_{(m,f,n)}(W,S)$ as a function of W and S.

The minimum of $AIC_m(W,S)$ (for the Class 1 *mQRS*) is found at $[(W_m^{\min}=20, S_m^{\min}=5]$, of $AIC_f(W,S)$ (for the Class 2 *fQRS*) is found at $[(W_f^{\min}=30, S_f^{\min}=10]$ and (for the Class 3 *noise*) is found at $[(W_n^{\min}=10, S_n^{\min}=3]$. We noticed that the AIC surfaces are quite flat around these points and that other model parameters can also suit our purpose with a slight improvement in the statistical indices shown in figure A2 and table 4.

Figure A2 shows the Sensitivity, Precision, and Specificity for the optimal number of splits (which corresponds to the minimum of the AIC $S_{(m,f,n)}^{\text{Optim}}$ for all W executed on the test set. The optimal points $[(W_{(m,f,n)}^{\text{Optim}}, S_{(m,f,n)}^{\text{Optim}}]$ are marked with a dark green dot.

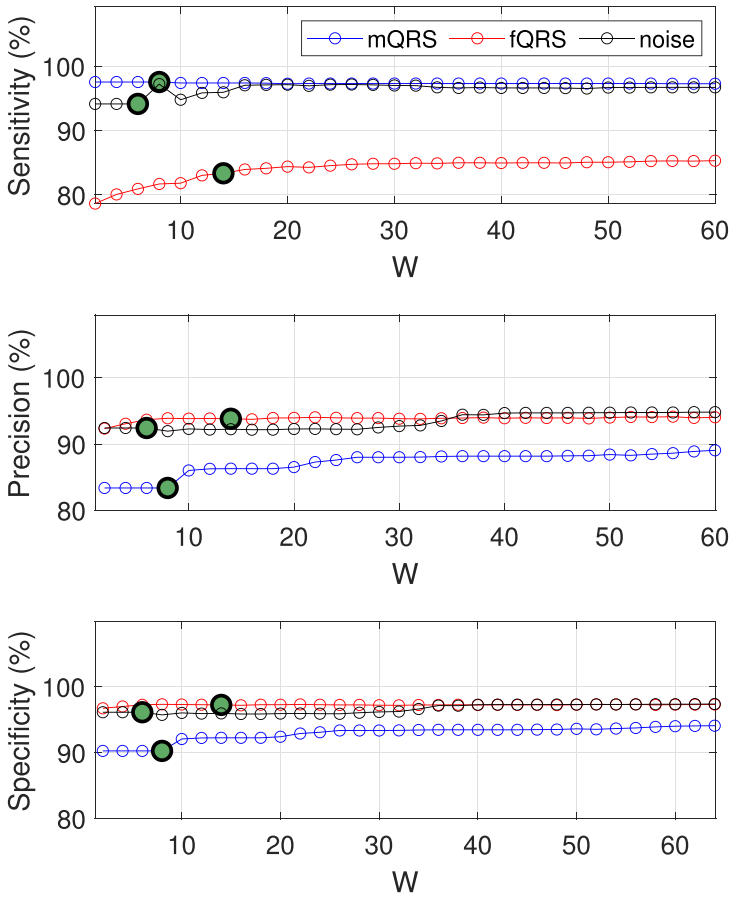

Figure A2. Sensitivity, Precision, and Specificity as a function of *W*.

Appendix B

Our approach is similar to the RUSBoost algorithm (Seiffert *et al* 2008) for its rather simple structure and lightness in memory usage and involves:

- Initializing the tree weights using a prior probability on the training set,
- Combining random undersampling and sorting selection of best samples with boosting,
- Use several boosting iteration with Gini impurity split criterion,
- Handling imbalanced and large dataset multiclass problems,
- Follows the core boosting logic: sequential learners, sample reweighting, ensemble voting,
- Fast ensemble grow stopping criteria to avoid overfitting.

The core differences between our algorithm and existing boosting ones are summarized in table B1.

The architecture of our classifier for this specific project is based on maximum $W = 60$ decision trees (boosting iterations), maximum $S = 20$ splits using $F = 3$ features and $Ns = 2300$ samples.

The weighted classification error of a weak learner on the current distribution of training data is denoted $\epsilon \in [0, 1]$ with $\epsilon < 0.5$ better and $\epsilon > 0.5$ worse than random guessing). The symbol ϵ refers to the weighted classification error of a weak learner on the current distribution of training data. We used the Gini index as tree split criterion (Strobl *et al* 2007).

As we have heavily unbalanced data (much more noise than *fQRS* and *mQRS* data samples) we used data rebalancing technique using random undersampling similar to the RUSboost algorithm (Seiffert *et al* 2008). The flow chart of our algorithm is outlined in figures B1 and B2. It includes an initialization phase (orange boxes) and, a training and best model selection phase (box with thick blue border).

Table B1. Differences between the our algorithm and classic AdaBoost.

Feature	AdaBoost	Our algorithm
Imbalance	AdaBoost struggles with class imbalance	Random undersampling is performed
Learner	Learns on full weighted data	Learns on a balanced undersampled subset
Multiclass	AdaBoost.M1 or SAMME required	The method supports multiclass by design
Stopping	Typically fixed rounds or early stopping	May assess performance per round and stop manually
Tree depth	Usually decision stumps	Allow *S* splits, making the weak learner stronger

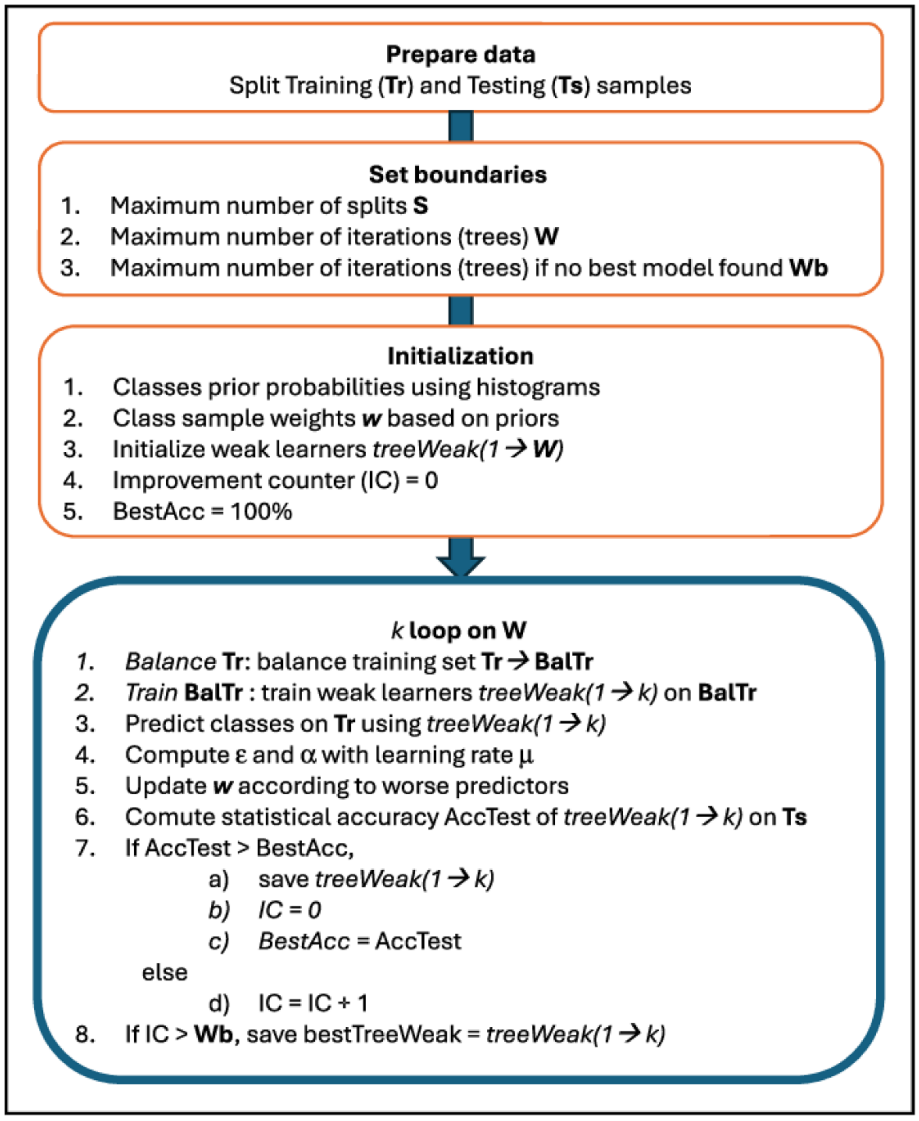

Figure B1. Flow chart of our algorithm.

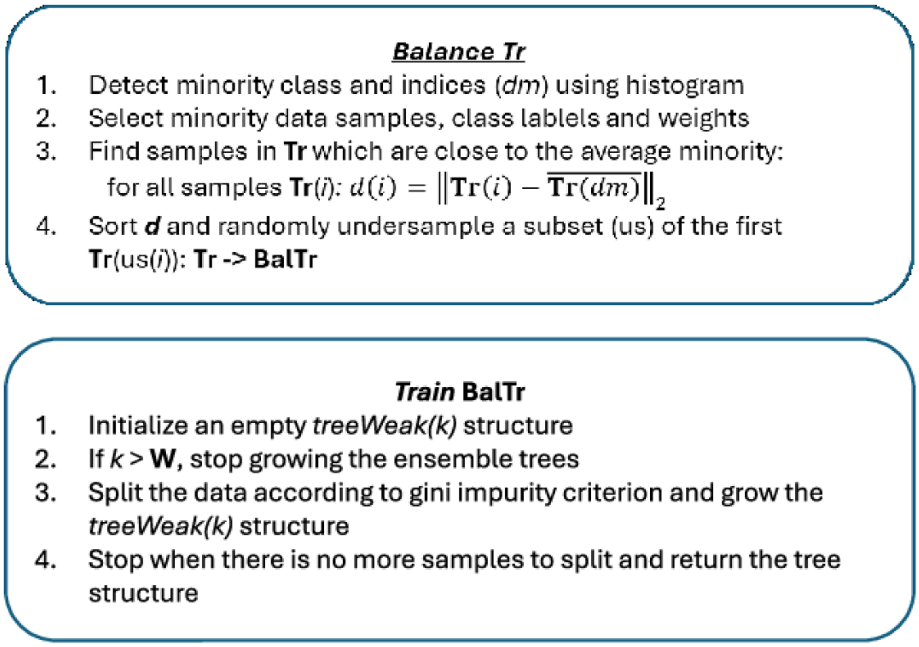

Figure B2. The two main processing blocks from figure B1. The symbol $\bar{\mathbf{x}}$ means the average value of $\mathbf{x}$ and the $||\mathbf{x}||_2$ means the L_2 norm of the vector $\mathbf{x}$.

Appendix C

Let T be the length of the observation sequence of fetus R-waves $fR(k)$. The table C1 introduces the symbols and their definitions in the Viterbi implementation.

Table C1. Definitions for our Viterbi algorithm.

Feature	Definition
S	The number M of possible heartbeat candidates
$fR(k)$	The detected heartbeat at time k from either of the two Branches
$x(k)$	The hidden state at time k which represent any state in S
$fRR(i,j)$	The IBIs $fR(i) - fR(j)$ following a Gaussian distribution $\mathcal{N}(fRR, \mu_{fRR}, \sigma_{fRR})$ with mean μ_{fRR} and variance σ^2_{fRR}
$a_{i,j}$	The transition probability of detecting a heartbeat in state $x(k) = j$ if $x(k-1) = i$ which obeys $\mathcal{N}(fRR(i,j), \mu_{fRR}, \sigma_{fRR})$
$B_j(fR(k))$	The emission probability of observation $fR(k)$ under state j: $B_j(fR(k)) = 1 - (1 - fSQI_{\mathrm{PT}}(k))(1 - fSQI_{\mathrm{C}}(k))$

We seek the most likely heartbeat sequence given observations $fR(0:T-1)$ where $(0:N)$ means the array of samples from 0 to N. We define the Viterbi score $V_k(j)$ up to time k ending in state j:

$$V_k(j) \triangleq \max_{x(0:k-2)} P(x(0:k-2), x(k) = j, fR(0:k-1)) \tag{C.1}$$

$$P(x(0:T-1)), fR(0:T-1)) = B_{x(0)}(fR(0)) \prod_{k=1}^{T-1} a_{x(k-1),x(k)} B_{x(k)}(fR(k)). \tag{C.2}$$

where $P(.)$ is the Markov probability. Because products of probabilities underflow quickly, we implement the recurrence in log-domain. Let $lV_t(j) = \log V_t(j)$, and denote la_{ij} and $lB_j(fR(k))$ the log-transition and log-emission. The recurrence for $k \geqslant 2$ is::

$$lV_0(j) \quad = lB_j(y_0), \tag{C.3}$$

$$lV_k(j) \quad = lB_j(y_k) + \max_{i \in \{1,2\}} \left\{ lV_{k-1}(i) + la_{ij} \right\}. \tag{C.4}$$

Algorithm 1. Viterbi Pseudo-Code.

Initialize empty lists for candidates and arrays V, psi.
for each new candidate j (time t_k, qualities $fSQI_{PT}, fSQI_C$) **do**
  $logB = log(1-(1-fSQI_P T)*(1-fSQI_C))$
  $V[j] = logB$
  $psi[j] = 0$
  for each previous candidate $i < j$ **do**
    $dt = t_j - t_i$
    if $dt \in [min_I BI, max_I BI]$ **then**
      $logTrans = -((dt-\mu_f RR)^2)/(2*\sigma_f RR^2) - \log(sqrt(2*\pi)*\sigma_f RR)$
      $score = V[i] + logTrans + logB$
    end if
    if $score > V[j]$ **then**
      $V[j] = score$
      $psi[j] = i$
    end if
  end for
  $Pickj* = \arg\max_j V[j]$
  Backtrack path using psi from $j*$ to start.
end for

To recover the event sequence that maximizes the Viterbi score, we store the time event indices backward in time in an array $\psi_k(j) = \arg\max_{i\in\{1,2\}}\{V_{k-1}(i)\,a_{ij}\}$ and backtrack: $\hat{x}(T-1) = \arg\max_j V_{T-1}(j)$ and for $k = T-2,\ldots,0$ set $\hat{x}(k) = \psi_{k+1}(\hat{x}(k+1))$. We finally store $\psi_k(j) = \arg\max_i\{lV_{k-1}(i) + \log a_{ij}\}$.

References

Abel J D K, Dhanalakshmi S and Kumar R 2023 A comprehensive survey on signal processing and machine learning techniques for non-invasive fetal ECG extraction *Multimedia Tools Appl.* **82** 1373–400

Ahmed M R, Newby S, Potluri P, Mirihanage W and Fernando A 2024 Emerging paradigms in fetal heart rate monitoring: evaluating the efficacy and application of innovative textile-based wearables *Sensors* **24** 6066

Akaike H 1998 Information theory and an extension of the maximum likelihood principle *Biogeochemistry* **1998** 199–213

Andreotti F, Behar J, Zaunseder S, Oster J and Clifford G D 2016 An open-source framework for stress-testing non-invasive foetal ECG extraction algorithms *Physiol. Meas.* **37** 627–48

Andreotti F, Graser F, Malberg H and Zaunseder S 2017 Non-invasive fetal ECG signal quality assessment for multichannel heart rate estimation *IEEE Trans. Biomed. Eng.* **64** 2793–802

Atkinson J, Hastie R, Walker S, Lindquist A and Tong S 2023 Telehealth in antenatal care: recent insights and advances *BMC Medicine* **21** 332

Barnhart H X, Haber M J and Lin L I 2007 An overview on assessing agreement with continuous measurements *J. Biopharmaceut. Stat.* **17** 529–69

Barnova K, Martinek R, Jaros R, Kahankova R, Matonia A, Jezewski M, Czabanski R, Horoba K and Jezewski J 2021 A novel algorithm based on ensemble empirical mode decomposition for non-invasive fetal ECG extraction *PLoS One* **16** e0256154

Bearak J, Popinchalk A, Ganatra B, Moller A B, Tunçalp Ozge, Beavin C, Kwok L and Alkema L 2020 Unintended pregnancy and abortion by income, region and the legal status of abortion: estimates from a comprehensive model for 1990–2019 *Lancet Glob. Health* **8** e1152–61

Behar J, Andreotti F, Zaunseder S, Li Q, Oster J and Clifford G D 2014 An ECG simulator for generating maternal-foetal activity mixtures on abdominal ECG recordings *Physiol. Meas.* **35** 1537

Benchekroun M, Chevallier B, Zalc V, Istrate D, Lenne D and Vera N 2023 The impact of missing data on heart rate variability features: A comparative study of interpolation methods for ambulatory health monitoring *IRBM* **44** 100776

Cavanaugh J E and Neath A A 2019 The akaike information criterion: Background, derivation, properties, application, interpretation and refinements *Wiley Interdiscip. Rev.- Comput. Stat.* **11** e1460

Chakraborty M and Das S 2012 Determination of signal to noise ratio of electrocardiograms filtered by band pass and Savitzky-Golay filters *Proc. Technol.* **4** 830–3

Chandraharan E, Ghi T, Fieni S and Jia Y-J 2023 Optimizing the management of acute, prolonged decelerations and fetal bradycardia based on the understanding of fetal pathophysiology *Am. J. Obstetr. Gynecol.* **228** 645–56

Chen L, Wu S and Zhou Z 2025 Fetal ECG signal extraction from maternal abdominal ECG signals using attention r2w-net *Sensors* **25** 601

Clifford G D, Silva I, Behar J and Moody G B 2014 Non-invasive fetal ECG analysis *Physiol. Meas.* **35** 1521

Darmawahyuni A, Tutuko B, Nurmaini S, Rachmatullah M N, Ardiansyah M, Firdaus F, Sapitri A I and Islami A 2023 Accurate fetal qrs-complex classification from abdominal electrocardiogram using deep learning *Int. J. Comput. Intell. Syst.* **16** 1–10

Darsana P and Kumar V N 2023 Extracting fetal ecg signals through a hybrid technique utilizing two wavelet-based denoising algorithms *IEEE Access* **11** 91696–708

de Plaza M A P, Yadav L and Kitson A 2023 Co-designing, measuring and optimizing innovations and solutions within complex adaptive health systems *Front. Health Serv.* **3** 1154614

Eenkhoorn C, Goos T G, Dankelman J, Franx A and Eggink A J 2024 Evaluation and patient experience of wireless noninvasive fetal heart rate monitoring devices *Acta Obs. Gynecol. Scand.* **103** 980–91

Fluss R, Faraggi D and Reiser B 2005 Estimation of the youden index and its associated cutoff point *Biometr. Z.* **47** 458–72

Freund Y and Schapire R E 1996 Experiments with a new boosting algorithm *Int. Conf. on Machine Learning* (available at: https://cseweb.ucsd.edu/~yfreund/papers/boostingexperiments.pdf)
Galli A, Peri E, Zhang Y, Vullings R, van der Ven M, Giorgi G, Ouzounov S, Harpe P J and Mischi M 2021 Dedicated algorithm for unobtrusive fetal heart rate monitoring using multiple dry electrodes *Sensors* **21** 4298
Gilgen-Ammann R, Schweizer T and Wyss T 2019 Rr interval signal quality of a heart rate monitor and an ECG holter at rest and during exercise *Eur. J. Appl. Physiol.* **119** 1525–32
Goldberger A L, Amaral L A, Glass L, Hausdorff J M, Ivanov P C, Mark R G, Mietus J E, Moody G B, Peng C K and Stanley H E 2000 Physiobank, physiotoolkit and physionet: components of a new research resource for complex physiologic signals *Circulation* **101** e215–20
Hernandez Engelhart C, Gundro Brurberg K, Aanstad K J, Pay A S D, Kaasen A, Blix E and Vanbelle S 2023 Reliability and agreement in intrapartum fetal heart rate monitoring interpretation: a systematic review *Acta Obs. Gynecol. Scand.* **102** 970–85
Huang Z, Yu J, Shan Y and Wang X 2024 A non-invasive fetal qrs complex detection method based on a multi-feature fusion neural network *Appl. Sci.* **14** 8987
Jamshidian-Tehrani F and Sameni R 2018 Fetal ECG extraction from time-varying and low-rank noninvasive maternal abdominal recordings *Phys. Meas.* **39** 125008
Jaros R, Martinek R, Kahankova R and Koziorek J 2019 Novel hybrid extraction systems for fetal heart rate variability monitoring based on non-invasive fetal electrocardiogram *IEEE Access* **7** 131758–84
Jaros R, Tomicova E and Martinek R 2024 Template subtraction based methods for non-invasive fetal electrocardiography extraction *Sci. Rep.* **14** 1–15
Jezewski J, Matonia A, Kupka T, Roj D and Czabanski R 2012 Determination of fetal heart rate from abdominal signals: evaluation of beat-to-beat accuracy in relation to the direct fetal electrocardiogram *Biomed. Eng. Appl. Basis Commun.* **57** 383–94
Jezewski J, Wrobel J, Matonia A, Horoba K, Martinek R, Kupka T and Jezewski M 2017 Is abdominal fetal electrocardiography an alternative to Doppler ultrasound for fhr variability evaluation? *Front. Physiol.* **8** 305
Joutsen A, Cömert A, Kaappa E, Vanhatalo K, Riistama J, Vehkaoja A and Eskola H 2024 ECG signal quality in intermittent long-term dry electrode recordings with controlled motion artifacts *Sci. Rep.* **14** 8882
Kahankova R, Martinek R, Jaros R, Behbehani K, Matonia A, Jezewski M and Behar J A 2019 A review of signal processing techniques for non-invasive fetal electrocardiography *IEEE Rev. Biomed. Eng.* **13** 51–73
Karlsson M, Hörnsten R, Rydberg A and Wiklund U 2012 Automatic filtering of outliers in rr intervals before analysis of heart rate variability in holter recordings: a comparison with carefully edited data *BioMed. Eng. Online* **11** 1–12
Keenan E, Karmakar C K and Palaniswami M 2018 The effects of asymmetric volume conductor modeling on non-invasive fetal ecg extraction *Physiol. Meas.* **39** 105013
Khandoker A, Ibrahim E, Oshio S and Kimura Y 2018 Validation of beat by beat fetal heart signals acquired from four-channel fetal phonocardiogram with fetal electrocardiogram in healthy late pregnancy *Sci. Rep.* **8** 1–11
Kruit H, Gissler M, Heinonen S and Rahkonen L 2022 Breaking the myth: the association between the increasing incidence of labour induction and the rate of caesarean delivery in finland - a nationwide medical birth register study *BMJ Open* **12** e060161
Kumar G, Duggal B, Singh J P and Shrivastava Y 2025 Efficacy of various dry electrode-based ecg sensors: A review *J. Biomed. Mater. Res. A* **113** e37845
Lampros T, Kalafatakis K, Giannakeas N, Tsipouras M G, Glavas E and Tzallas A T 2023 An optimized hybrid methodology for non-invasive fetal electrocardiogram signal extraction and monitoring *Array* **19** 100302
Lee K J and Lee B 2022 End-to-end deep learning architecture for separating maternal and fetal ecgs using w-net *IEEE Access* **10** 39782–8
Lempersz C, Noben L, Vries B D, Laar J O V, Westerhuis M E and Oei S G 2020 The noninvasive fetal electrocardiogram during labor: a review of the literature *Obs. Gynecol. Surv.* **75** 369–80
Li X, Wan J and Peng X 2025 Review of non-invasive fetal electrocardiography monitoring techniques *Sensors* **25** 1412
Likitalo S, Pakarinen A and Axelin A 2025 Integrating remote monitoring into the pregnancy care perspectives of pregnant women and healthcare professionals *CIN - Comput. Inf. Nursing* **43** e01255
Liu B, Thilaganathan B and Bhide A 2023 Effectiveness of ambulatory non-invasive fetal electrocardiography: impact of maternal and fetal characteristics *Acta Obs. Gynecol. Scand.* **102** 577–84
Matonia A, Jezewski J, Kupka T, Horoba K, Wrobel J and Gacek A 2006 The influence of coincidence of fetal and maternal qrs complexes on fetal heart rate reliability *Med. Biol. Eng. Comput.* **44** 393–403
Matonia A, Jezewski J, Kupka T, Jezewski M, Horoba K, Wrobel J, Czabanski R and Kahankowa R 2020 Fetal electrocardiograms, direct and abdominal with reference heartbeat annotations *Sci. Data* **7** 1–14
Murphy K P 2012 *Machine Learning* (MIT Press)
Neilson J P 2015 Fetal electrocardiogram (ECG) for fetal monitoring during labour *Cochrane Database Syst. Rev.* **2015** CD000116
Nichting T J, Frenken M W, van der Woude D A, van Oostrum N H, de Vet C M, van Willigen B G, van Laar J O, van der Ven M and Oei S G 2021 Non-invasive fetal electrocardiography, electrohysterography and speckle-tracking echocardiography in the second trimester: study protocol of a longitudinal prospective cohort study (beats-study) *BMC Pregn. Childbirth* **21** 1–11
O'Brien E, Rauf Z, Alfirevic Z and Lavender T 2013 Women's experiences of outpatient induction of labour with remote continuous monitoring *Midwifery* **29** 325–31
Orvas I, Radu A, Galli A, Neacsu A and Peri E 2025 A complex unet approach for non-invasive fetal ECG extraction using single-channel dry textile electrodes (arXiv:2506.22457)
Pan J and Tompkins W J 1985 A real-time qrs detection algorithm *IEEE Trans. Biomed. Eng.* **32** 230–6
Pawłowski R, Al-Ani F, Samjeed A, Buszko K and Khandoker A 2025 Investigating asymmetry in fetal and maternal heart rate accelerations and decelerations *Sci. Rep.* **15** 1–18
Pimentel M A, Santos M D, Springer D B and Clifford G D 2014 Hidden semi-markov model-based heartbeat detection using multimodal data and signal quality indices *Comput. Cardiol.* **41** 553–6
Ponsiglione A M, Cosentino C, Cesarelli G, Amato F and Romano M 2021 A comprehensive review of techniques for processing and analyzing fetal heart rate signals *Sensors* **21** 6136
Rahman S, Karmakar C, Natgunanathan I, Yearwood J and Palaniswami M 2022 Robustness of electrocardiogram signal quality indices *J. R. Soc. Interface* **19** 20220012
Rooijakkers M J, Song S, Rabotti C, Oei S G, Bergmans J W, Cantatore E and Mischi M 2014 Influence of electrode placement on signal quality for ambulatory pregnancy monitoring *Comput. Math. Methods Med.* **2014** 960980
Ross M G 2011 Labor and fetal heart rate decelerations: relation to fetal metabolic acidosis *Clin. Obs. Gynecol.* **54** 74–82

Rowe S, Karkhaneh Z, MacDonald I, Chambers T, Amjad S, Osornio-Vargas A, Chari R, Kumar M and Ospina M B 2020 Systematic review of the measurement properties of indices of prenatal care utilization *BMC Pregn. Childbirth* **20** 171
Saarikko J, Niela-Vilen H, Ekholm E, Hamari L, Azimi I, Liljeberg P, Rahmani A M, Löyttyniemi E and Axelin A 2020 Continuous 7-month internet of things-based monitoring of health parameters of pregnant and postpartum women: prospective observational feasibility study *JMIR Form. Res.* **4** e12417
Sameni R, Clifford G D, Jutten C and Shamsollahi M B 2007 Multichannel ECG and noise modeling: Application to maternal and fetal ECG signals *EURASIP J. Adv. Signal Process.* **2007** 1–14
Samuel B and Hota M K 2024 A novel hybrid method for calculating the fetal heart rate from the non-invasive abdominal electrocardiogram signal *Biomed. Signal Process. Control* **94** 106277
Searle A and Kirkup L 2000 A direct comparison of wet, dry and insulating bioelectric recording electrodes *Physiol. Meas.* **21** 271–83
Seiffert C, Khoshgoftaar T M, Hulse J V and Napolitano A 2008 Rusboost: improving classification performance when training data is skewed *Proc. - Int. Conf. on Pattern Recognition* (https://doi.org/10.1109/ICPR.2008.4761297)
Serrano L P *et alet al* 2023 Benefits and challenges of remote patient monitoring as perceived by health care practitioners: a systematic review *Permanente J.* **27** 100–11
Shaffer F and Ginsberg J P 2017 An overview of heart rate variability metrics and norms *Front. Public Health* **5** 258
Shi X, Yamamoto K, Ohtsuki T, Matsui Y and Owada K 2023 Unsupervised learning-based non-invasive fetal ECG muti-level signal quality assessment *Bioengineering* **10** 66
Silva I, Lee J and Mark R G 2012 Signal quality estimation with multichannel adaptive filtering in intensive care settings *IEEE Trans. Biomed. Eng.* **59** 2476–85
Skála T, Vícha M, Rada M, Vácha J, Flašík J and Táborský M 2022 Feasibility of evaluation of polar h10 chest-belt ECG in patients with a broad range of heart conditions *Cor et Vasa* **64** 411–22
Smith V, Arunthavanathan S, Nair A, Ansermet D, Costa F D S and Wallace E M 2018 A systematic review of cardiac time intervals utilising non-invasive fetal electrocardiogram in normal fetuses *BMC Pregn. Childbirth* **18** 1–15
Souriau R, Fontecave-Jallon J and Rivet B 2023 Fetal heart rate monitoring by fusion of estimations from two modalities: a modified viterbi's algorithm *Biomed. Signal Process. Control* **80** 104405
Stricker K, Radan A P and Surbek D 2025 Continuous remote home monitoring solutions for mother and fetus: a scoping review *Eur. J. Obs. Gynecol. Reprod. Biol.* **305** 170–7
Strobl C, Boulesteix A L and Augustin T 2007 Unbiased split selection for classification trees based on the gini index *Comput. Stat. Data Anal.* **52** 483–501
Tijus C, Lee P-L, Yang C-F, Chang C-Y, Uddin R and Koo I 2024 Real-time remote patient monitoring: a review of biosensors integrated with multi-hop iot systems via cloud connectivity *Appl. Sci.* **14** 1876
Uv J J, Maleckar M M and Arevalo H 2025 Impact of vernix caseosa distribution on non-invasive fetal ECG morphology: a computational study *IEEE Trans. Biomed. Eng.* **72** 846–55
Viterbi A J 1967 Error bounds for convolutional codes and an asymptotically optimum decoding algorithm *IEEE Trans. Inf. Theory* **13** 260–9
Wahbah M, Zitouni M S, Sakaji R A, Funamoto K, Widatalla N, Krishnan A, Kimura Y and Khandoker A H 2024 A deep learning framework for noninvasive fetal ECG signal extraction *Front. Physiol.* **15** 1329313
Walter J R, Xu S, Stringer J S and Rogers J A 2022 The future of remote monitoring for pregnancy *Bridge* **52** 16–24
Wang L and Zhao C 2022 Fetal ECG extraction using a novel polynomial network and non-linear bss method *Electron. Lett.* **58** 750–2
Yerande V, Bhole K, Sonawane D N and Patil C Y 2022 A hybrid technique for non-invasive fetal ECG extraction and heart rate estimation from the mother's abdomen signal *Proc. - 2022 Int. Conf. on Artificial Intelligence of Things, ICAIoT 2022* (https://doi.org/10.1109/ICAIoT57170.2022.10121878)
Yuan L, Yuan Y, Zhou Z, Bai Y and Wu S 2019 A fetal ECG monitoring system based on the android smartphone *Sensors* **19** 446
Zander V, Gustafsson C, Stridsberg S L and Borg J 2023 Implementation of welfare technology: a systematic review of barriers and facilitators *Disab. Rehabil. Assist. Technol.* **18** 913–28
Zhao Z, Anand R and Wang M 2019 Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform *Proc. - 2019 IEEE Int. Conf. on Data Science and Advanced Analytics, DSAA 2019* pp 442–52
Zhou P, Schwerin B and So S 2024 U-net based fetal r-peak prediction from abdominal ECG signals *2024 9th Int. Conf. on Signal and Image Processing, ICSIP 2024* pp 121–5
Zhou P, So S and Schwerin B 2025 Robust u-nets for fetal r-peak identification in electrocardiography *Algorithms* **18** 487
Ziani S 2024 Enhancing fetal electrocardiogram classification: a hybrid approach incorporating multimodal data fusion and advanced deep learning models *Multimedia Tools Appl.* **83** 55011–51
Zizzo A R, Hvidman L, Salvig J D, Holst L, Kyng M and Petersen O B 2022 Home management by remote self-monitoring in intermediate- and high-risk pregnancies: a retrospective study of 400 consecutive women *Acta Obs. Gynecol. Scand.* **101** 135–44