Extracting fetal heart beats from maternal abdominal recordings: selection of the optimal principal components

Costanzo Di Maria[1,2]**, Chengyu Liu**[2]**, Dingchang Zheng**[2]**, Alan Murray**[2] **and Philip Langley**[3]

[1] Regional Medical Physics Department, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
[2] Institute of Cellular Medicine, Medical School, Newcastle University, Newcastle upon Tyne, UK
[3] School of Engineering, University of Hull, Hull, UK

E-mail: costanzo.dimaria@nuth.nhs.uk

Abstract

This study presents a systematic comparison of different approaches to the automated selection of the principal components (PC) which optimise the detection of maternal and fetal heart beats from non-invasive maternal abdominal recordings.

A public database of 75 4-channel non-invasive maternal abdominal recordings was used for training the algorithm. Four methods were developed and assessed to determine the optimal PC: (1) power spectral distribution, (2) root mean square, (3) sample entropy, and (4) QRS template. The sensitivity of the performance of the algorithm to large-amplitude noise removal (by wavelet de-noising) and maternal beat cancellation methods were also assessed. The accuracy of maternal and fetal beat detection was assessed against reference annotations and quantified using the detection accuracy score F1 [2*PPV*Se / (PPV + Se)], sensitivity (Se), and positive predictive value (PPV). The best performing implementation was assessed on a test dataset of 100 recordings and the agreement between the computed and the reference fetal heart rate (fHR) and fetal RR (fRR) time series quantified.

The best performance for detecting maternal beats (F1 99.3%, Se 99.0%, PPV 99.7%) was obtained when using the QRS template method to select the optimal maternal PC and applying wavelet de-noising. The best performance for detecting fetal beats (F1 89.8%, Se 89.3%, PPV 90.5%) was obtained when the optimal fetal PC was selected using the sample entropy method and utilising a fixed-length time window for the cancellation of the maternal beats. The performance on the test dataset was 142.7 beats^2/min^2 for fHR and

19.9 ms for fRR, ranking respectively 14 and 17 (out of 29) when compared to the other algorithms presented at the Physionet Challenge 2013.

(Some figures may appear in colour only in the online journal)

1. Introduction

Monitoring of fetal health throughout pregnancy is paramount in order to elicit prompt diagnosis of possible clinical complications and inform appropriate intervention (Kennedy 1998). The two systems most commonly used in current clinical practice are the cardiotocograph (CTG) and the fetal scalp electrode. A CTG (Spencer 1992) consists of two probes: a pressure transducer for monitoring uterine contractions, and an echo transducer for monitoring fetal heart rate (fHR). The information provided by this technology is limited to the fHR and a complete fetal ECG (fECG) signal is not available. Moreover, it is relatively expensive, needs a trained nurse for the appropriate placement of the probes, and is extremely sensitive to maternal and fetal movement. Compared to the CTG, the use of an electrode placed on the scalp of the fetus (Bartlett *et al* 1991; Bartlett *et al* 1992) provides more information as it collects the fECG signal from which not only the fHR, but also other important clinical parameters can be derived. However, the fetal scalp can only be used during labour after rupture of the membranes and is therefore highly invasive and carries additional risk of infection.

The limitations of these technologies along with the need to develop alternative methods to improve maternal and fetal care have been recognised for some time (Jenkins 1989). Non-invasive electrical recordings can be performed by placing electrodes on the abdomen of the mother. This approach, like the CTG, is non-invasive and, like the fetal scalp electrode, collects the fECG and hence offers the prospects of providing more comprehensive clinical parameters. Compared to current clinical standard, this method can also facilitate long-term monitoring. The reliability of cardiac parameters extracted from abdominal fetal ECG (aECG) recordings and their excellent agreement with the values obtained from a scalp electrode signal has been recently demonstrated (Clifford *et al* 2011). However, acquisition of clean signals from the maternal abdomen is not simple and the challenges in terms of signal processing to extract the fECG from aECG recordings are significant. In particular, these signals are contaminated with noise originated by different sources, including movement of the mother and fetus, interference from the mains power line, and other physiological components such as maternal ECG (mECG) and maternal muscular activity (Sameni and Clifford 2010).

Different signal processing techniques have been investigated so far for the extraction of the fetal ECG from non-invasive recordings. One of the most common is Independent Component Analysis (ICA) (Sameni *et al* 2006). However, ICA used alone has practical limitations due to the significant noise usually affecting these signals, and the small amplitude of the fECG as compared to the other physiological components. In order to overcome such limitations, ICA has been utilised in conjunction with wavelet decomposition (Azzerboni *et al* 2005); or as part of a 2-stage Blind Adaptive Filtering approach (Graupe *et al* 2005). An alternative approach for separating the different components of aECG signals is by Principal Component Analysis (PCA). PCA is based on obtaining statistically uncorrelated components, which do not necessarily represent independent physiological sources. This method was used, for example, in the algorithm developed by Martens *et al* (2007). Nevertheless, the reliable extraction of

the fetal ECG from abdominal recordings still remains an open challenge exemplified by the Physionet Challenge 2013 (Physionet Challlenge 2013 website) which promoted the development of advanced signal processing techniques to improve the state of the art in this field. A variety of methods were proposed, including subspace decomposition and reconstruction, adaptive filtering and averaging, wavelet de-noising, matched filtering, Christov's beat detection, entropy, RS slope, expectation weighting, and echo state recurrent neural network (Silva *et al* 2013). Most of the algorithms were structured in four steps, respectively for (1) preprocessing, (2) maternal heart beat detection, (3) maternal heart beat cancellation, and (4) fetal heart beat detection. With respect to the steps for detection of maternal and fetal heart beats, one possible approach is to detect the beats from all the available channels before proceeding to further analysis, as done for example by Lipponen and Tarvainen (2013). Another possible approach requires the selection of an optimal channel on which the heart beat detection will be performed, as done for example by Andreotti and colleagues (Andreotti *et al* 2013). In the latter case, different strategies are possible to select this optimal channel. Similarly, for a PCA based approach such as that used in our study (Di Maria *et al* 2013), it is important to determine the algorithms for selecting the principal components which optimise subsequent heart beat detection.

The aim of this study were: (i) to compare different methods for automatically selecting the principal components which optimised the detection of maternal and fetal beats, (ii) to assess the importance of large-amplitude noise removal and several approaches to maternal ECG cancellation, and (iii) to compare the performance of the algorithm to that of other algorithms presented in the Physionet Challenge 2013.

2. Methods

2.1. Dataset

The dataset Set-A from the Computing in Cardiology Physionet Challenge 2013 was utilised as the training set for developing and assessing the different processing tools. This dataset contains 75 4-channel non-invasive abdominal recordings, each one lasting 60 s. For each recording, a reference fetal beat time series (fR-peaks) was also provided. For our study, to enable the accuracy of automatic detection of maternal beats to be quantified, the maternal beats (mR-peaks) were manually annotated by the first author. The annotated mR-peaks were selected on the channel which visually showed the most clear maternal ECG after basic 3–100 Hz filtering. All recordings were visually inspected to assess recording quality. Manual annotation of the maternal beats was not possible across the entire recording in 4 recordings (records a29, a38, a54, and a56), so these were removed from the analysis. Hence, 71 recordings (SET-1) were available to test the accuracy of maternal beat detection. Visual inspection showed that, out of these 71 recordings, 59 did not contain large-amplitude noise (SET-1a), and 12 did (SET-1b). An additional 5 recordings had incomplete annotations of the fetal beats (records a33, a47, a52, a71, a74) and these were removed from the stages of the analysis following maternal beat detection. Hence, 66 recordings (SET-2) were available to test the accuracy of fetal beat detection.

The performance of the best implementation was finally assessed on the test dataset Set-B, also from the Computing in Cardiology Physionet Challenge 2013, for which reference fetal R-peaks were not available to us. This dataset consisted of 100 4-channel recordings, each one lasting 60 s. Both Set-A and Set-B are available online through the Physionet website (Goldberg *et al* 2000) and described in more detail in an editorial (Clifford *et al* 2014).
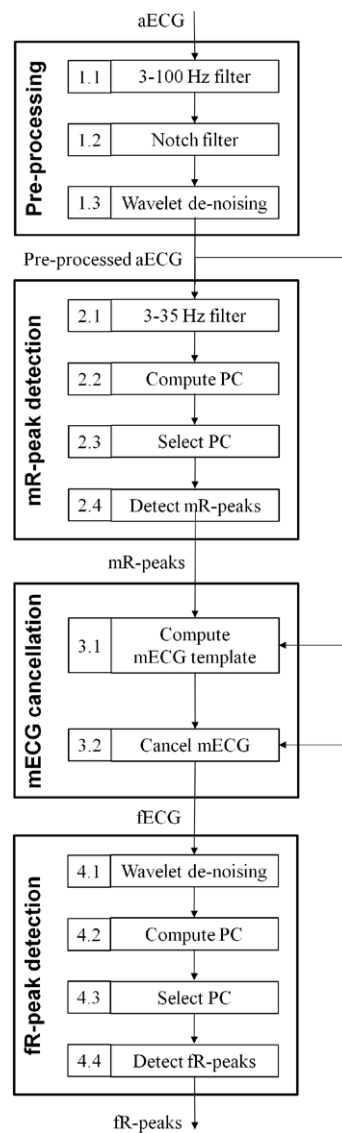
Figure 1. Flow diagram showing the different steps of the algorithm.

2.2. Detection of maternal and fetal beats using PCA (initial algorithm)

The initial algorithm for detection of maternal and fetal beats was based on PCA of the multi-lead recordings and using always the first principal component (PC) for detecting either the maternal or fetal beats, i.e. without any optimisation of the selected PC. This section describes the initial algorithm and section 2.3 presents the additional processing tools for selecting the PC which were assessed in this work leading to the final implementation.

The general structure of the algorithm consisted of four main steps (figure 1)**:** **(**1) pre-processing of the raw aECG signals for noise reduction; (2) detection of the maternal beats; (3) cancellation of the maternal beats; (4) detection of the fetal beats. The specific implementation of each of these in the initial algorithm is described below.
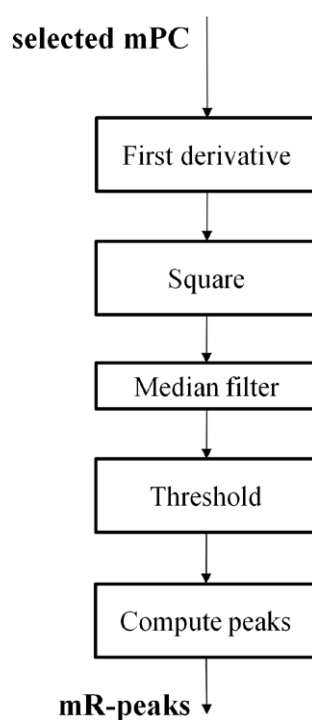
Figure 2. Flow chart illustrating the steps involved in the method used for detecting the maternal R-peaks (mR-peaks), starting from the selected maternal principal component (mPC).

Pre-processing. Low and high frequency noise, mostly caused by movement artefacts, respiration, and electrical interference (Onaral *et al* 1984), was removed by a 3–100 Hz band-pass filter (figure 1, step 1.1). Power line noise was suppressed by an appropriate notch filter (step 1.2). Large–amplitude spikes, for example due to sudden movement artefact, were removed by wavelet de-noising (step 1.3). For this purpose, a 1-level wavelet decomposition was applied to each of the four channels using a bi-orthogonal mother wavelet. A threshold of 10 times the root mean square value was defined for the approximation; a threshold of 0.15 times the root mean square value was defined for the detail. The approximation and detail components were set to zero when they were larger than their respective threshold.

Detection of maternal beats. To emphasise the features of the maternal QRS (mQRS) complex, a 3–35 Hz band-pass filter was applied (step 2.1) and the 4 maternal principal components (mPC) computed (step 2.2). The first principal component was selected for maternal beat detection (mPC_1) since this component was expected to contain the large maternal QRS complexes (step 2.3). Beat detection (figure 2) used thresholding of the square of the first derivative of mPC_1 to 1.5 times its root mean square value after applying a 40-point median filter, and selecting the peaks above this threshold (step 2.4).

Cancellation of maternal beats. An average maternal beat was calculated separately for each channel (step 3.1) using the automatically detected mR-peak times as reference points and considering the heart beat to be contained in an RR-adjusted-length time window including 30% of the previous and next RR time interval. The maternal beats were cancelled by aligning the average beat with each maternal beat according to the lag which gave the highest correlation. Then the template was scaled both for width and amplitude and subtracted from

the original signal leaving the fetal ECG (step 3.2). This process was carried out separately for each of the four original channels and for the two halves of the signal (first and last 30 s) in order to account for space and time variability of the ECG waveform.

Detection of fetal beats. Wavelet de-noising was used to reduce artefacts of mECG cancellation (step 4.1). For this purpose, a threshold was fixed based on the root mean square value of the *n*th-level detail signal. The coefficient for the root mean square value used at leach level 1–10 was respectively: 3.5, 8, 0, 0, 0, 80, 100, 85, 80, 65. The signal of the *n*th-level detail was set to zero when below its respective threshold before reconstruction. The four fetal principal components (fPC) were computed (step 4.2). The first principal component was selected for fetal beat detection because this component was expected to contain the (relatively) large fetal ECG QRS complex (step 4.3). Fetal beats were detected using a local peak detection algorithm.

2.3. Algorithms for selecting the optimal principal component for maternal and fetal peak detection

In the above, for both maternal and fetal beat detection, the first principal component was used based on the expectation that this would contain the large ventricular activity. However, this expectation might not always be met, particularly in noisy recordings, and the current study assessed algorithms for selecting the principal component which provided the best performance for maternal and fetal beat detection. Furthermore, optimal cancellation of maternal beats is likely to affect the performance of the algorithm so this study also assessed the sensitivity of the algorithm to different implementations of maternal beat cancellation.

2.3.1. Selection of the optimal principal component. Four approaches for selecting the optimal PC were implemented: (1) power spectral distribution (psd) (for maternal beat detection only); (2) root mean square (rms); (3) sample entropy (sampen); and (4) QRS template.

The aim with each of these approaches was to select the PC with the clearest and most enhanced QRS complexes, as this would facilitate the process of detecting the peaks of the R waves. The rationale for these approaches follows.

Power spectral distribution. It was expected that most of the power of the QRS is in the band 5–35 Hz. A psd score was calculated for each principal component as the ratio of the power in the band 5–35 Hz (as representative of the QRS complex) to the power in the band 0–5 Hz (as representative of the low-frequency component of the ECG) as defined in equation (1.1). This was done for six 10 s non-overlapping portions of the principal component and the average value used as the final psd score (equation (1.2)). The PC with the largest psd score was selected for beat detection.

$$\text{psd\_ratio}_i = \frac{\text{PSD}_{5-35\,\text{Hz}}}{\text{PSD}_{0-5\,\text{Hz}}} \tag{1.1}$$

$$\text{psd\_score} = \frac{1}{6}\sum_{i=1}^{6}\text{psd\_ratio}_i \tag{1.2}$$

Root mean square. This method was based on the idea of quantifying the amplitude of the QRS complexes in each principal component. An rms score was calculated for each principal component as follows. A threshold *th* was defined as the root mean square value of the squared signal x^2. An rms score was calculated as the ratio of the sum of the amplitudes of the signal above the threshold (related to the R waves) to the sum of the amplitudes of the signal below

the threshold (related to the baseline level of the signal) as defined in equation (2.1). This was done for six 10 s non-overlapping portions of the principal component and the average value used as the final rms score (equation (2.2)). The PC with the largest score was selected for subsequent beat detection.

$$\text{rms\_ratio}_i = \sum_{k, x_k^2 \geq \text{th}} x_k^2 \bigg/ \sum_{k, x_k^2 < \text{th}} x_k^2 \tag{2.1}$$

$$\text{rms\_score} = \frac{1}{6} \sum_{i=1}^{6} \text{rms\_ratio}_i \tag{2.2}$$

Sample entropy. Sample entropy quantifies the regularity of a signal. A very regular signal will have a very low value of sample entropy whereas a signal containing random noise will have a higher value. Therefore, this method for the selection of the mPC is based on the assumption that a principal component with clear ECG and little noise will have lower values of sampen. For the calculation of sample entropy, a value of $m = 3$ was used for the embedding dimension, and a value of $r = 0.15$ times the standard deviation was used for the tolerance. Implementation details for the calculation of sample entropy are provided in Richman and Moorman (2000). A sample entropy value was calculated for each of the six 10 s non-overlapping portions of a principal component. For each principal component, a sampen score was obtained as the average of these six values. The PC with the lowest sampen score was chosen to be the optimal one for subsequent beat detection.

QRS template. The QRS template approach firstly produced a smoothed version of each PC. An ECG-like template (represented by a 100 ms triangular wave) was passed along the PC and the correlation coefficient between the PC and the triangular wave calculated at each step. A threshold *th* was defined as the root mean square value of the correlation signal *C*. A template score was calculated as the ratio of the mean value of the correlation signal above the threshold (related to the amplitude of the R waves) and the threshold value itself (as a measure of the baseline level of the signal) as defined in equation (3.1). The final template score was given by the average of the six ratios calculated in each of the 10 s portions of the signal (equation (3.2)). The PC with the largest value for this score was chosen as the optimal one for subsequent beat detection.

$$\text{template\_ratio}_i = \frac{\sum\limits_{i, x_i^2 \geq \text{th}} C}{\text{th}} \tag{3.1}$$

$$\text{template\_score} = \frac{1}{6} \sum_{i=1}^{6} \text{template\_ratio}_i \tag{3.2}$$

2.3.2. Cancellation of maternal beats. Maternal beat cancellation was based on the generation of an average beat followed by the subtraction of this from each beat in the waveform. Two important considerations for average beat generation were (i) the size of the window assumed to contain the beat, and (ii) the number of beats used to generate the average beat. The initial approach described above in section 2.2 was to use an adjusted-length time window (30% of the previous and following RR time) and 30 beats to calculate the average one. To assess the sensitivity of the algorithm to specific implementation of this general framework for maternal beat cancellation, the initial approach was compared to one using a fixed-length time window for the interval considered to contain the maternal beat and also to different numbers

of beats used for the average beat generation. For the fixed-length time window approach, the maternal ECG complex was considered to be in a window containing the 250 ms before and the 450 ms after the mR-peak. The number of beats used to generate the average beat ranged from 5 to 30 in steps of 5.

2.4. Assessment

Maternal beat detection. Assessment of the methods for selecting the optimal maternal PC was in terms of accuracy in detection of the maternal beats using the manual annotation as the reference, on the basis that the optimal PC would provide the greatest accuracy. The effect of wavelet de-noising (figure 1, step 1.3) on the accuracy of maternal beat detection was assessed by comparing the performance of the algorithm with and without this noise reduction step. The implementation providing the most accurate maternal beat detection was retained for subsequent assessment of the fetal beat detection.

Fetal beat detection. Assessment of the methods for selection of the optimal fetal PC was done in terms of accuracy in the detection of fetal beats against the reference annotations. Sensitivity of the algorithm to specific implementations of maternal beat cancellation (specifically, fixed or adjusted time window and variable number of beats used to generate the average beat) was assessed also in terms of accuracy of fetal beat detection.

2.5. Scoring

For evaluation of the performance of the algorithm on the training set in detecting the maternal and fetal R-peaks as compared to the reference values, primarily a detection accuracy score (F1, Behar *et al* 2013) was utilised (equation (4.1)) to provide a metric which balances sensitivity and positive predictive value.

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{4.1}$$

Where TP is the number of correctly identified R-peaks, FP is the number of falsely identified R-peaks, and FN is the number of missed R-peaks. A detected R-peak was labelled as TP if within 100 ms of a reference R-peak. If no corresponding reference R-peak was found in this time window, then the detected R-peak was labelled as FP. If no corresponding detected R-peak was found for a reference R-peak in the same time window, this generated a FN. As further assessment metrics, also the sensitivity (Se) and the positive predictive value (PPV) were considered.

$$Se = \frac{TP}{TP + FN} \tag{4.2}$$

$$PPV = \frac{TP}{TP + FP} \tag{4.3}$$

The total F1 score, Se, and PPV for the entire dataset were calculated as the average of the individual values obtained for each record.

The performance on the test set (for the best performing implementation identified from the training set) was quantified using: (1) the mean square error between the reference and the computed fetal heart rate (fHR) time series measured in beats^2/min^2, and (2) the average root square error between the reference and computed fetal RR (fRR) time series measured in milliseconds (ms) (Clifford *et al* 2014). Lower values of these metrics indicate better performance

Table 1(*a*). Performance for detecting the maternal beats on the entire dataset SET-1. The best performance results are shown in bold. WD is wavelet de-noising; F1 is the detection accuracy score; Se is sensitivity; and PPV is positive predictive value.

Method	F1 (%)		Se (%)		PPV (%)	
	with WD	no WD	with WD	no WD	with WD	no WD
mPC_1	97.4	93.5	98.5	94.4	96.8	93.0
psd	78.2	80.2	78.9	80.8	79.5	81.5
rms	96.3	93.6	96.3	93.5	97.1	94.0
sampen	98.0	95.2	97.6	94.8	99.0	95.8
template	**99.3**	97.9	**99.0**	97.6	**99.7**	98.2

Table 1(*b*). Performance for detecting the maternal beats on the subset without records containing large-amplitude noise (SET-1a). The best performance results are shown in bold. WD is wavelet de-noising; F1 is the detection accuracy score; Se is sensitivity; and PPV is positive predictive value.

Method	F1 (%)		Se (%)		PPV (%)	
	with WD	no WD	with WD	no WD	with WD	no WD
mPC_1	98.3	98.3	99.1	99.2	97.8	97.7
psd	79.3	79.3	80.0	80.1	79.8	79.8
rms	99.1	99.1	99.0	99.0	99.2	99.2
sampan	**99.4**	**99.4**	99.0	99.0	**99.8**	**99.8**
template	**99.4**	**99.4**	**99.1**	**99.1**	**99.8**	**99.8**

Table 1(c). Performance for detecting the maternal beats on the subset with only records containing large-amplitude noise (SET-1b). The best performance results are shown in bold. WD is wavelet de-noising; F1 is the detection accuracy score; Se is sensitivity; and PPV is positive predictive value.

Method	F1 (%)		Se (%)		PPV (%)	
	with WD	no WD	with WD	no WD	with WD	no WD
mPC_1	92.9	70.3	95.2	71.3	92.0	69.7
psd	73.0	84.7	73.1	84.5	78.0	89.5
rms	83.0	66.5	82.7	66.4	86.8	68.2
sampen	91.2	74.5	90.9	74.4	95.0	76.2
template	**99.0**	90.5	**99.0**	90.6	**99.1**	90.5

(less error). This allowed for direct comparison of our results with the performance of other published work from the Physionet Challenge 2013. The fHR and fRR scores were calculated using the online scoring tool available in the relevant PhysioNetWorks webpage.

3. Results

3.1. Selection of the optimal mPC and effect of wavelet denoising

Table 1 shows the performance for the three indices F1, sensitivity, and positive predictive value of the various methods used for selecting the optimal mPC, with and without wavelet denoising. Table 1(*a*) shows the results for the entire SET-1. Table 1(*b*) presents the results for the subset of 59 records which did not contain large-amplitude noise (SET-1a). Table 1(*c*)

gives the results for the 12 records containing large-amplitude noise (SET-1 b). The implementation which provided the best performance was the one using the template method for the selection of the optimal maternal PC along with wavelet de-noising. The enhanced performance of this implementation over the others in noisy recordings is clear since it achieved a detection accuracy score F1 of 99%, whereas selecting always the first PC (mPC$_1$) achieved a score of less than 93% (table 1(*c*)). This implementation gave F1 99.3%, Se 99.0%, and PPV 99.7% on the entire SET-1 for detection of the maternal R peaks (table 1(*a*)).

3.2. Selection of the optimal fPC and maternal beat cancellation

Figure 3 shows the performance in terms of F1 (figure 3(*a*)), Se (figure 3(*b*)) and PPV (figure 3(*c*)) for different methods of selecting the fetal PC. The figure also shows the effect on performance of the different methods for maternal beat cancellation considering fixed and adjusted-length time windows and different numbers of beats used for generating the average maternal ECG beat. Since the psd method gave very poor performance for the maternal beat detection it was not further considered for the selection of the fetal PC. The best performance was obtained using the sample entropy method for selecting the fPC and a fixed-length time window for generating the average maternal beat from 30 heart cycles. This final implementation gave accuracy in detecting the fetal heart beats of F1 89.8%, Se 89.3%, and PPV 90.5%. This compared favourably to the performance of the initial algorithm which was F1 75.1%, Se 73.0%, and PPV 78.2%.

Figure 4 shows the application of the final algorithm to record a05 of Set-A through the different processing steps. Figure 4(*a*) shows the four channels of the raw abdominal ECG recording; figure 4(*b*) shows the same signals after pre-processing and including wavelet de-noising; figure 4(*c*) shows the four channels after performing maternal beat cancellation; figure 4(*d*) shows the fetal PC selected for final detection of the fetal heart beats. The computed beats are indicated by red arrows and the reference ones by black full circles.

3.3. Results on the test set and comparison with published algorithms

The final best performing implementation identified above was then assessed on the test dataset Set-B. This gave a result of 142.7 beats^2/min^2 for the fHR time series and 19.9 ms for the fRR time series. Table 2 compares the performance of this implementation of our algorithm with the published results of the 28 other participants to the Physionet Challenge 2013 for fHR (event 4) and fRR (event 5) scores on the open test dataset Set-B. The dataset used to score the other events was not publicly available.

4. Discussion

This study presented a systematic comparison of different methods for the automated selection of the principal components which optimise the detection of maternal and fetal beats in the context of fetal ECG analysis from non-invasive maternal abdominal recordings.

Five methods were evaluated for the selection of the optimal maternal PC: (1) first principal component, (2) power spectral distribution, (3) root mean square, (4) sample entropy, and (5) template. The results showed that the template algorithm performed best overall considering the entire SET-1 (table 1(*a*)). This template method can be considered as a simplified version of the approach presented by Andreotti *et al* (2013) which assessed the agreement between the channels and an average beat obtained using Gaussian kernels. Sample entropy performed equally
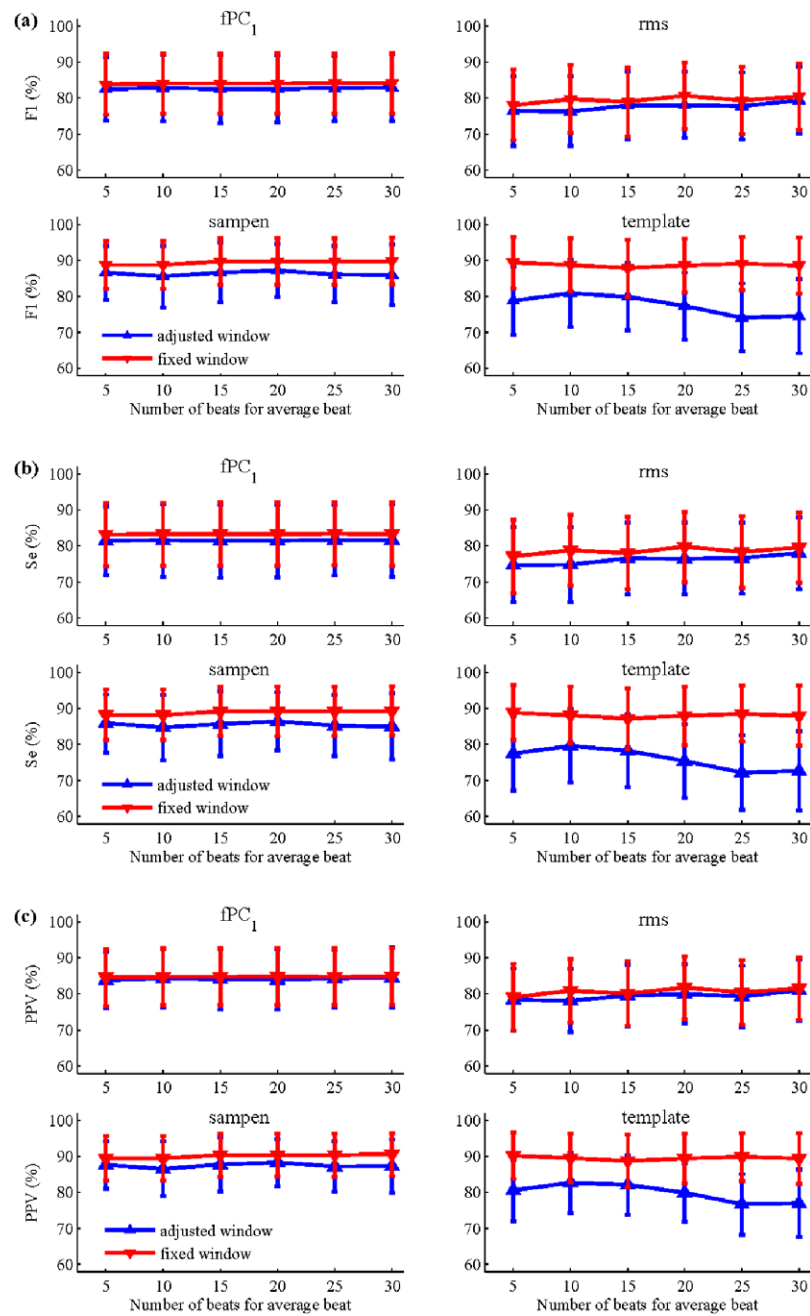
Figure 3. Performance for the detection of the fetal R peaks quantified as (*a*) detection accuracy score F1, (*b*) sensitivity Se, and (*c*) positive predictive value PPV. Results are given for each of the four methods used for the selection of the optimal fetal PC, and comparing the two different approaches to maternal ECG cancellation, i.e. adjusted-length time window (blue line) versus fixed-length time window (red line). Results are also shown separately for the different values of number of beats used for generating the maternal average beat. The error bars represent the mean value (triangle) obtained on SET-2 and the standard error of the mean.
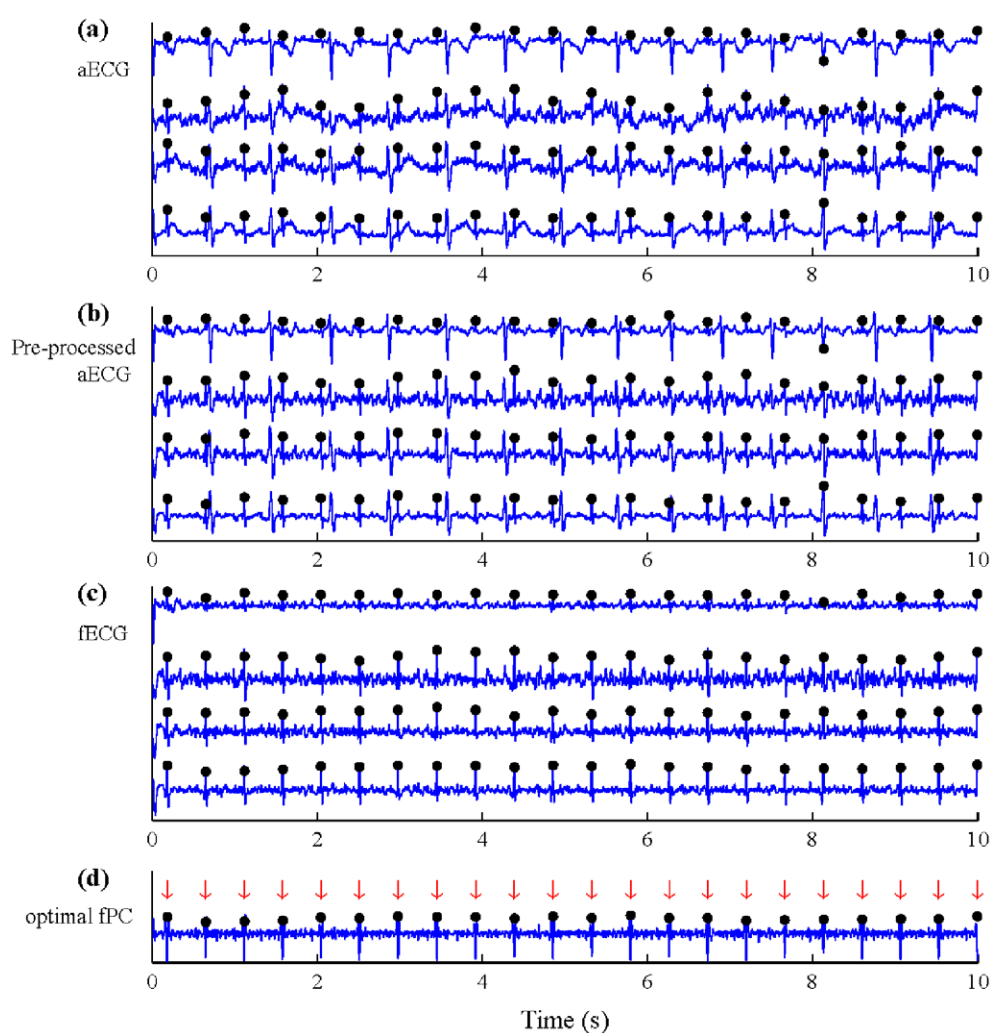
Figure 4. Record a05 at different steps of the processing: (*a*) raw 4-channel signals, (*b*) after pre-processing—including wavelet de-noising, (*c*) after maternal beat cancellation, and (*d*) optimal fetal principal component selected for fetal beat detection. An excerpt of 10 s is shown. Computed (red arrows) and reference (black full circles) fetal R peaks are indicated.

well with signals not affected by large-amplitude noise (SET-1a, table 1(*b*)). The improvement in using the template method was even clearer with signals containing large-amplitude noise, giving an F1 of 99.0% on SET-1 b (table 1(*c*)). In this case the sample entropy performed worse than using the first principal component, F1 91.2% and 92.9% respectively. This result is not surprising considering that the wavelet de-noising method actually modifies the signals setting to zero the portions containing large-amplitude noise. This will have inevitably affected the regularity of the signal as quantified by sample entropy. The value of performing wavelet de-noising before computing the principal components was shown by the consistently better performance of each method as compared to the case without wavelet de-noising (table 1(*a*)).

Table 2. Performance of this algorithm on the test dataset Set-B in comparison to other results from the Physionet Challenge 2013 as published in the proceedings of Computing in Cardiology 2013.

Publication	fHR score (beats^2/min^2) [rank out of 29]	fRR score (ms) [rank out of 29]
Andreotti *et al* (2013)	18.1 [1]	4.4 [1]
Lipponen and Tarvainen (2013)	28.9 [2]	4.8 [3]
Behar *et al* (2013)	29.6 [3]	4.7 [2]
Varanini *et al* (2013)	34.0 [4]	5.1 [4]
Haghpanahi and Borkholder (2013)	50.1 [5]	9.1 [9]
Xu-Wilson *et al* (2013)	52.5 [6]	10.6 [11]
Ghaffari *et al* (2013)	63.8 [7]	11.2 [14]
Lukoševičius and Marozas (2013)	66.3 [8]	8.2 [7]
Kropf *et al* (2013)	82.4 [9]	7.4 [6]
Podziemski and Gieraltowski (2013)	118.2 [10]	10.7 [12]
Maier and Dickhaus (2013)	118.4 [11]	9.4 [10]
Rodrigues (2013)	124.8 [12]	14.4 [15]
Di Marco *et al* (2013)	135.2 [13]	7.1 [5]
Starc (2013)	181.0 [15]	10.9 [13]
Razavipour *et al* (2013)	210.1 [16]	21.2 [19]
Di Maria *et al* (2013)	223.2 [17]	19.3 [16]
Kuzilek and Lhotska (2013)	249.8 [18]	22.0 [20]
Perlman *et al* (2013)	262.1 [19]	27.8 [23]
Liu and Li (2013)	264.9 [20]	9.0 [8]
Fatemi *et al* (2013)	274.3 [21]	32.1 [24]
Christov *et al* (2013)	285.1 [22]	20.0 [18]
Petrolis and Krisciukaitis (2013)	341.5 [23]	32.8 [25]
Almeida *et al* (2013)	521.4 [24]	33.0 [26]
Dessì *et al* (2013)	639.5 [25]	23.8 [21]
Plešinger *et al* (2013)	688.5 [26]	26.8 [22]
Akhbari *et al* (2013)	1326.2 [27]	45.1 [27]
Niknazar *et al* (2013)	1514.6 [28]	57.0 [28]
Llamedo *et al* (2013)	4714.6 [29]	121.6 [29]
This algorithm	**142.7 [14]**	**19.9 [17]**

Only the psd method contradicted this result and actually performed better in the absence of wavelet de-noising. This is probably because the large-amplitude noise contained power in the 3–35 Hz band contributing to the selection of a principal component which was not ideal for detecting the maternal beats. When the signals did not contain large-amplitude noise, all methods performed equally either with or without wavelet de-noising (table 1(*b*)). This gives confidence that the wavelet de-noising algorithm did not modify the signals when the quality of the recording was good. When signals contained large-amplitude noise, the use of wavelet de-noising greatly improved the performance (table 1(*c*)). In this case, the basic approach using the first principal component improved by 22.6% in F1 when using wavelet de-noising (from 70.3% to 92.9%). The best performing template method improved by 8.5% (from 90.5% to 99.0%). The use of wavelet de-noising allowed for an initially noisy channel to be kept for subsequent analysis on the assumption that portions of this noisy channel may still contain good quality signal which could help the subsequent detection of the fetal beats. This is an alternative approach to excluding noisy channels from the analysis (Liu and Li 2013).

Four methods were evaluated for the selection of the optimal fetal PC: (1) first principal component, (2) root mean square, (3) sample entropy, and (4) template. The method based on power spectral distribution performed poorly in the detection of maternal beats and therefore was not considered for the selection of the fetal PC. The best performing method was sample entropy which gave a mean F1 of 89.8% when using a fixed-length window for maternal beat

cancellation and 30 heart cycles for generating the maternal average beat (figure 3(*a*)). The template method also performed well (with fixed-length window), but not as well as sample entropy in this case. This is because the template method assumes that a clear QRS complex should be present which although almost always true for maternal beats, this is not the case for fetal beats due to their relatively low amplitude. The sample entropy was able to identify and quantify the regularity given by the fetal heart activity even in the absence of clear QRS complexes in the recordings. The use of a fixed-length time window for generating the average maternal beat always outperformed the adjusted-length window (figure 3). This is because the use of a 700 ms window is likely to contain almost the entire heart cycle whereas the adjusted-length window only contained about 60% of it and this may not suffice to include the entire PQRST complex in all cases. This approach led to non accurate cancellation in the case of slow heart rates in which the QT interval was relatively long and was a limitation of our initial algorithm. In addition, the adjusted-length window approach assumed a linear relationship between the duration of the PQRST complex and the RR time. However, a previous study (Malik *et al* 2002) showed that the relationship between the QT interval (here approximated to the PQRST length) and the RR time is not necessarily linear. Furthermore, even when this relationship is linear it has different slope from subject to subject. The algorithm was relatively insensitive to the number of beats utilised to generate the average maternal beat (figure 3).

After this assessment, the best performing implementation of the algorithm used the template method for selection of the optimal maternal PC and the sample entropy for the optimal fetal PC; it utilised the wavelet de-noising algorithm and a fixed-length time window for cancellation of the maternal beats (with 30 maternal beats used to generate the average beat). Compared to the initial algorithm (section 2.2) this implementation improved the performance for the ultimate goal of accurately detecting the fetal beats with F1 89.8% vs 75.1%, Se 89.3% vs 73.0% and PPV 90.5% vs 78.2%. Algorithm thresholds and parameters were set empirically and there was no systematic optimisation of these values in the present study, so this is a potential area of improvement. The step for reducing large-amplitude noise was considered also by other authors using different techniques. For example, Varanini and colleagues (Varanini *et al* 2013) used an approach based on selective median filtering. Other algorithms which performed particularly well in the challenge used more complex and adaptive methods for the cancellation of the maternal ECG such us Kalman smoother (Andreotti *et al* 2013), augmented principal component regression (Lipponen and Tarvainen 2013), and singular value decomposition (Varanini *et al* 2013). The best performing algorithms undertook a common four step approach but also included a post-processing stage to ensure parameters were within expected physiological ranges (Andreotti *et al* 2013). In addition, the best performing algorithms generated additional waveforms based on channel differences or ICA (Andreotti *et al* 2013; Lipponen and Tarvainen 2013), with the benefit of noise cancellation or feature enhancement. Adaptive processes, such as maternal template adaption, were common features of the best performing algorithms (Andreotti *et al* 2013; Haghpanahi and Borkholder 2013). Also, performance was enhanced when several algorithms were combined and a quality measure used to define the optimal algorithm for a specific recording (Behar *et al* 2013; Xu-Wilson *et al* 2013). For example, Behar *et al* (2013) used several algorithms for mECG extraction on each recording, and then from the resulting fHR time series chose the smoothest one. In addition, a post-processing stage, not considered by our algorithm, identified outliers in the fRR or fHR time series such that values lying outside of expected physiological range were constrained (Andreotti *et al* 2013). As a common decomposition technique PCA was the basis of many of the algorithms submitted to the Challenge (Di Maria *et al* 2013; Petrolis and Krisciukaitis 2013). However, without optimisation of all the processing steps, PCA alone could not match the performance of the highest ranking algorithms.

The final implementation presented in this work ranked 14 (fHR score) and 17 (fRR score) out of 29 when compared to the other algorithms presented at the Physionet Challenge 2013.

5. Conclusion

The results of this study suggest the template and the sample entropy methods as promising approaches to selecting the optimal PCA channel for maternal and fetal heart beat detection. The value of using wavelet de-noising has also been demonstrated along with the sensitivity of the performance of the algorithm to the method used for cancellation of the maternal beats.

References

Akhbari M, Niknazar M, Jutten C, Shamsollahi M B and Rivet B 2013 Fetal electrocardiogram R-peak detection using robust tensor decomposition and extended Kalman filtering *Comput. Cardiol.* **40** 189–92

Almeida R, Gonçalves H, Rocha A P and Bernardes J 2013 A wavelet-based method for assessing fetal cardiac rhythms from abdominal ECGs *Comput. Cardiol.* **40** 289–92

Andreotti F, Riedl M, Himmelsbach T, Wedekind D, Zaunseder S, Wessel N and Melberg H 2013 Maternal signal estimation by Kalman filtering and template adaptation for fetal heart rate extraction *Comput. Cardiol.* **40** 193–6

Azzerboni B, La Foresta F, Mammone N and Morabito F C 2005 A new approach based on wavelet-ICA algorithms for fetal electrocardiogram extraction *Proc. ESANN'2005* pp 193–8 ISBN 2-930307-05-6;

Bartlett M L R, Murray A and Dunlop W 1991 Properties of fetal heart beat intervals during labour *J. Biomed. Eng.* **13** 169–72

Bartlett M L R, Murray A and Dunlop W 1992 Is fetal heart rate monitoring sufficiently sensitive to detect changes during labour? *J. Biomed. Eng.* **14** 431–4

Behar J, Oster J and Clifford G D 2013 Non-invasive FECG extraction from a set of abdominal sensors *Comput. Cardiol.* **40** 297–300

Christov I, Simova I and Abächerli R 2013 Cancellation of the maternal and extraction of the fetal ECG in noninvasive recordings *Comput. Cardiol.* **40** 153–6

Clifford G D, Sameni R, Ward J, Robinson J and Wolfberg A J 2011 Clinically accurate fetal ECG parameters acquired from maternal abdominal sensors *Am. J. Obstet. Gynecol.* **205** 47.e1-5

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

Dessì A, Pani D and Raffo L 2013 Identification of fetal QRS complexes in low density non-invasive biopotential recordings *Comput. Cardiol.* **40** 321–4

Di Marco L Y, Marzo A and Frangi A 2013 Multichannel foetal heartbeat detection by combining source cancellation with expectation-weighted estimation of fiducial points *Comput. Cardiol.* **40** 329–32

Di Maria C, Wenfeng D, Bojarnejad M, Pan F, King S, Zheng D, Murray A and Langley P 2013 An algorithm for the analysis of foetal ECG from 4-channel non-invasive abdominal recordings *Comput. Cardiol.* **40** 305–8

Fatemi M, Niknazar M and Sameni R 2013 A robust framework for noninvasive extraction of fetal electrocardiogram signals *Comput. Cardiol.* **40** 201–4

Ghaffari A, Atyabi S A, Mollakazemi M J, Niknazar M, Niknami M and Soleimani A 2013 PhysioNet/CinC Challenge 2013: a novel noninvasive technique to recognize fetal QRS complexes from noninvasive fetal electrocardiogram signals *Comput. Cardiol.* **40** 293–6

Goldberger A L, Amaral L A N, Glass L, Hausdorff J M, Ivanov P Ch, 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

Graupe D, Zhong Y and Graupe MH 2005 Extraction of fetal ECG from maternal ECG early in pregnancy *IJBEM* **7** 166–8

Haghpanahi M and Borkholder D A 2013 Fetal ECG extraction from abdominal recordings using array signal processing *Comput. Cardiol.* **40** 173–6

Jenkins H M L 1989 Thirty years of electronic intrapartum fetal heart rate monitoring: discussion paper *J. R. Soc. Med.* **82** 210–4

Kennedy R G 1998 Electronic fetal heart rate monitoring: retrospective reflections on a twentieth-century technology *J. R. Soc. Med.* **91** 244–50

Kropf M, Modre-Osprian R, Schreier G and Hayn D 2013 A robust algorithm for fetal QRS detection using non-invasive maternal abdomenal ECGs *Comput. Cardiol.* **40** 313–6

Kuzilek J and Lhotska L 2013 Advanced signal processing techniques for fetal ECG analysis *Comput. Cardiol.* **40** 177–80

Lipponen J A and Tarvainen M P 2013 Advanced maternal ECG removal and noise reduction for application of fetal QRS detection *Comput. Cardiol.* **40** 161–4

Liu C and Li P 2013 Systematic methods for fetal electrocardiographic analysis: determining the fetal heart rate, RR interval and QT interval *Comput. Cardiol.* **40** 309–12

Llamedo M, Martín-Yebra A, Laguna P and Martínez J P 2013 Noninvasive fetal ECG estimation based on linear transformations *Comput. Cardiol.* **40** 285–8

Lukoševičius M and Marozas V 2013 Noninvasive fetal QRS detection using echo state network *Comput. Cardiol.* **40** 205–8

Maier C and Dickhaus H 2013 Fetal QRS detection and RR interval measurement in noninvasively registered abdominal ECGs *Comput. Cardiol.* **40** 165–8

Malik M, Färbom P, Batchvarov V, Hnatkova K and Camm A J 2002 Relation between QT and RR intervals is highly individual among healthy subjects: implications for heart rate correction of the QT interval *Heart* **87** 220–8

Martens S M M, Rabotti C, Mischi M and Sluijter R J 2007 A robust fetal ECG detection method for abdominal recordings *Physiol. Meas.* **28** 373–88

Niknazar M, Rivet B and Jutten C 2013 Fetal QRS complex detection based on three-way tensor decomposition *Comput. Cardiol.* **40** 185–8

Onaral B, Sun H H and Schwan H P 1984 Electrical properties of bioelectrodes *IEEE Trans. Biomed. Eng.* **31** 827–32

Perlman O, Katz A and Zigel Y 2013 Noninvasive fetal QRS detection using a linear combination of abdomen ECG signals *Comput. Cardiol.* **40** 169–72

Petrolis R and Krisciukaitis A 2013 Multi stage principal component analysis based method for detection of fetal heart beats in abdominal ECGs *Comput. Cardiol.* **40** 301–4

Physionet Challenge 2013 website: www.physionet.org/challenge/2013/

Plešinger F, Jurák P and Halámek J 2013 Extracting the R-wave position from an FECG record using recognition of multi-channel shapes *Comput. Cardiol.* **40** 157–60

Podziemski P and Gierałtowski J 2013 Fetal heart rate discovery: algorithm for detection of fetal heart rate from noisy, noninvasive fetal ECG recordings *Comput. Cardiol.* **40** 333–6

Razavipour F, Haghpanahi M and Sameni R 2013 Fetal QRS complex detection using semi-blind source separation framework *Comput. Cardiol.* **40** 181–4

Richman J S and Moorman J R 2000 Physiological time-series analysis using approximate entropy and sample entropy *Am. J. Physiol. Heart Circ. Physiol.* **278** H2039-49

Rodrigues R 2013 Fetal ECG detection in abdominal recordings: a method for QRS location *Comput. Cardiol.* **40** 325–8

Sameni R and Clifford G D 2010 A review of fetal ECG signal processing issues and promising directions *Open Pacing Electrophysiol. Ther. J.* **3** 4–20

Sameni R, Jutten C and Shamsollahi M B 2006 What ICA provides for ECG processing: application to noninvasive fetal ECG extraction *IEEE ISSPIT* **2006** 656–61

Silva I, Behar J, Sameni R, Zhu T, Oster J, Clifford G D and Moody G B 2013 Noninvasive fetal ECG: the PhysioNet/computing in cardiology challenge 2013 *Comput. Cardiol.* **40** 149–52

Spencer J A 1992 Role of cardiotocography *Br. J. Hosp. Med.* **48** 115–8

Starc V 2013 Non-invasive fetal multilead RR interval determination from maternal abdominal recordings: the Physionet/CinC challenge 2013 *Comput. Cardiol.* **40** 317–20

Varanini M, Tartarisco G, Billeci L, Macerata A, Pioggia G and Balocchi R 2013 A multi-step approach for non-invasive fetal ECG analysis. *Comput. Cardiol.* **40** 281–4

Xu-Wilson M, Carlson E, Cheng L and Vairavan S 2013 Spatial filtering and adaptive rule based fetal heart rate extraction from abdominal fetal ECG recordings *Comput. Cardiol.* **40** 197–200