Intelligent Heartsound Diagnostics on a Cellphone using a Hands-free Kit

T. Chen [a], K. Kuan [a], L. Celi [a,b], G. D. Clifford [a,b,c]

[a] Massachusetts Institute of Technology, [b] Harvard Medical School, [c] University of Oxford

[a] 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[b] 221 Longwood Avenue, Boston, MA 02115, USA
[c] Institute of Biomedical Engineering, Department of Engineering Science, Oxford, UK

Abstract

In resource-constrained environments, supply chains for consumables, repairs and calibration of diagnostic equipment are generally poor. To obviate this issue, we propose the use of widely available hardware with a strong supply chain: a cellphone with a hands-free kit. In particular, we focus on the use of the audio channel to determine heart rate (HR) and heart rate variability (HRV) in order to provide a first level screening system for infection. This article presents preliminary work performed on a gold standard database and a cellphone platform. Results indicate that HR and HRV can be accurately assessed from acoustic recordings of heart sounds using only a cellphone and hands-free kit. Heart sound analysis software, which can run on a standard cellphone in real time, has been developed that detects S1 heart sounds with a sensitivity of 92.1% and a positive predictivity of 88.4%. Evaluation of data recorded from cellphones demonstrates that the low-frequency response (<100 Hz) is key to the success of heart sound analysis on cellphones. Noise rejection is also shown to be important.

Introduction and Background

There is a wide rural-urban divide in health care delivery, especially in developing nations. Medical specialists in these countries are scarce and are often only found in the cities. For people living in remote or resource-poor locations, travel to see these specialists can deprive them of a whole day's income. For many rural clinics, the time it takes to send information to the nearest physician and receive a diagnosis and advice can take weeks. As a result, diagnosis and treatment are often delayed and patient follow-up is difficult when a long journey or wait time is involved, resulting in higher mortality and costs than are necessary. Although training programs exist to increase the numbers of community health workers, such programs are not scalable and sustainable, requiring constant resources, the effectiveness of which is reduced as the knowledge radiates out from the centers of training.

Maternal and childhood mortality is a particularly pressing issue. Each year, over half a million women die from pregnancy or childbirth (Richards 2009). Furthermore, women in least developed countries are 300 times more likely to die in childbirth. With proper prenatal care and routine screening, mothers can learn to take proper safety measures during pregnancy, including preparations for a high-risk delivery if necessary. Through simple monitoring of the mother and fetus (i.e. measuring fetal heartbeat and respiration), a healthcare worker would be able to check on the heart condition and general growth of the baby and any infections of the mother. Following childbirth, infections in the young children (such as TB) require detection.

One promising method for such screening is through fetal and pediatric heart rate (HR) analysis. Changes in heart rate variability (HRV) have been shown to be linked to infection (Kovatchev *et al* 2003, Blad *et al* 2008, Frasch *et al* 2009). However, to-date all analyses of fetal and pediatric heart rate variability has been via ultrasound or electrocardiogram (ECG), with the exception of a prototype computer-base system in India (Mittra 2009), which uses high-end microphones to subtract ambient noise and render a heart rate. However, Mittra gives no details of the heart rate extraction or its accuracy on the small number of mothers tested. Moreover, a large amount of equipment is required to perform the screening.

In the area of mobile diagnostics using cellular technologies there have been several recent developments. For example, Jin *et al* (2009) have developed a system to record ECGs via a cellphone. Tan and Masek (2009) have developed a system to interface with Doppler devices for fetal ultrasound assessment. Black *et al* (2009) have created a low cost pulse oximeter attached to a cellphone to try to distinguish pneumonia from other febrile illnesses. However, all these systems require a reliable supply chain infrastructure to deliver the diagnostic peripherals, maintenance and supplies. Furthermore, perhaps with the exception of the pulse oximeter, expert knowledge and training is usually required to use the equipment.

In this article we present an analysis of heart-rate extraction from off-line acoustic data and compare them to that extracted from a clean electrocardiogram. We then present results from a preliminary cellphone-based system which performs the same operation. Finally we describe a telemedicine framework for collecting expert-labeled data, an essential requirement for training our system.

Methods

A database of cardiac acoustic and electrocardiogram (ECG) data, and cellphone recordings were analyzed in Matlab to determine the accuracy of audio beat detection.

Data Sources. ECG and heart sounds were previously recorded with a Master Elite Plus Welch Allyn Meditron electronic stethoscope for approximately 30 seconds for each subject (Syed, 2003). The stethoscope had a frequency response of 20 Hz to 20 kHz, and files were stored as WAV files without compression. Recordings were performed with the Bell setting, which applied a bandpass filter from 20 Hz to 420 Hz.

Out of 123 recordings, 27 clean ECG (with associated audio recordings of undetermined quality) of healthy adults with no noted heart abnormalities were chosen for analysis. Data was collected from subjects in a supine position from the left lower sternal border (tricuspid area), as shown in Figure 1. Recording examples are given in Figure 2.
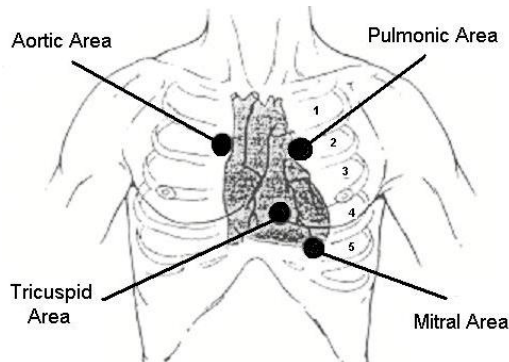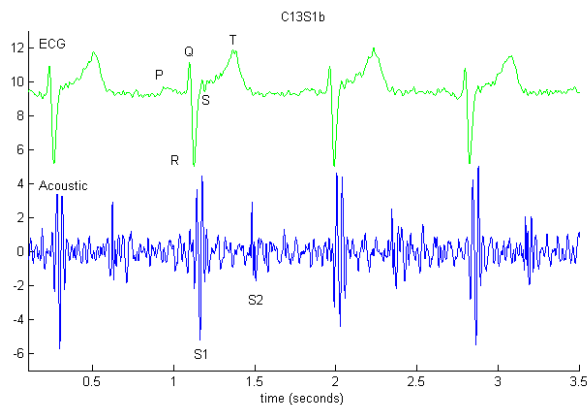

Figure 1: Auscultation sites

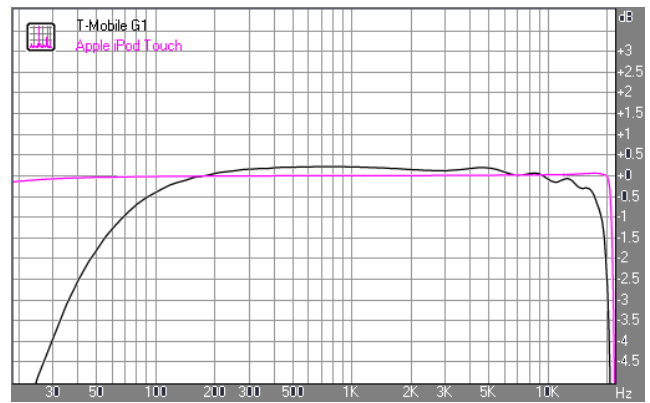**Figure 2: ECG and associated heart sounds**

Figure 3: Sound card frequency response (dB/Hz) comparison between HTC G1 (black line) and iPhone (magenta line). Adapted from GSMArena 2009.

Recordings were also performed using an Apple iPhone 3GS and a HTC G1 phone with a hands-free kit from a young, healthy female subject at rest from the pulmonic area. Data was recorded in *m4a* format and converted to WAV for processing. All data was sampled at 44.1 kHz with 16-bit quantization. The frequency response of the two test phones is given in Figure 3. The rationale for choosing these two phones is that they provide a high quality starting point to being a feasibility study.

Preprocessing. ECG and acoustic data were downsampled to 500 Hz and passed through 100-point FIR bandpass filters. The low and high cutoff frequencies for the ECG filter were 2 Hz and 30 Hz, respectively, to reduce high frequency noise and baseline wander. For data collecting from the electronic stethoscope, acoustic recordings were filtered to preserve the frequency range from 5 Hz to 70 Hz, which was determined empirically.

Event Detection and HR Estimation. HR estimation was performed by detection the onset of S1 and S2 sounds and, for data recorded using the electronic stethoscope, comparing them to QRS detection in ECG. The peak detection algorithm was based on an ECG QRS detection method (Pan and Tompkins 1985, Hamilton and Tompkins 1986). The energy of the heart sound signal was quantified by differentiating, squaring, and integrating over a fixed-length window, which was empirically determined to be 26 milliseconds for QRS detection. The resulting integrated quantity peaked during high energy areas, specifically during the QRS complex, shown in Figure 4. The R peaks were identified by thresholding the integrated quantity and searching for the location of the peaks above the threshold.

Event detection for heart sounds was performed using the same procedure to derive the integrated quantity, with an integration window of 58 milliseconds. Local maxima of the integrated value were identified to demarcate S1 and S2 sounds.

To distinguish S1 from S2 sounds, the time interval (SS interval) between detected local maxima were computed. A distribution of these SS intervals was created for each

record, as shown in Figure 5, resulting in two Gaussian-like clusters of S1-S2 and S2-S1 intervals. Each SS interval was then classified as either S1-S2 or S2-S1 by its proximity to the cluster centers. Since S1-S2 intervals are typically shorter than S2-S1 intervals, the cluster of shorter SS intervals was designated as S1-S2, and the longer SS intervals as S2-S1. The start and end points of each S1-S2 interval was identified as S1 and S2 sounds, as in Figure 6.
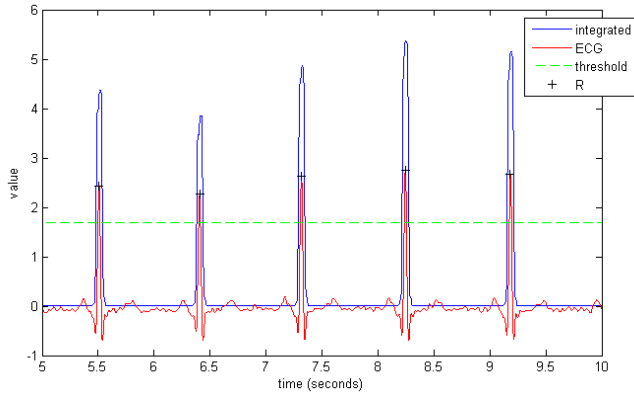

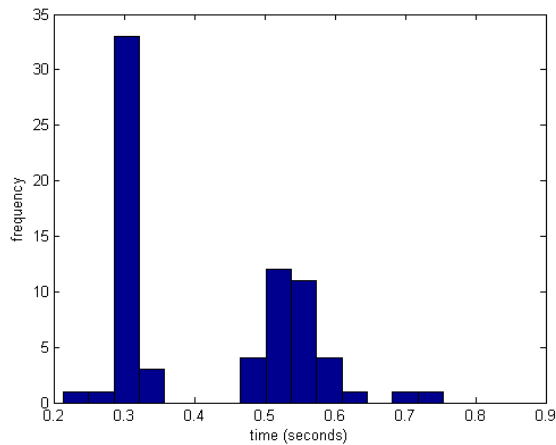**Figure 4: ECG R peak detection using integrated signal**

Figure 5: Empirical distribution of SS intervals for entire record of patient depicted in Figure 4
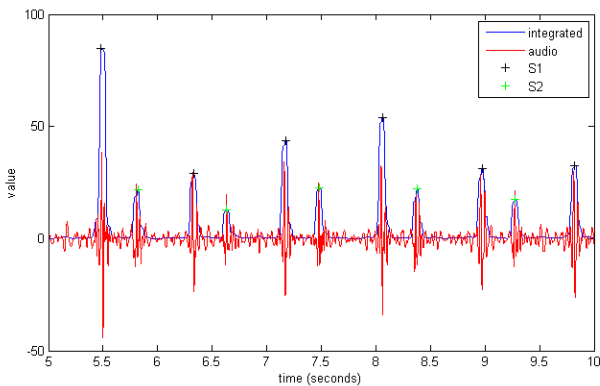

Figure 6: S1 and S2 detection using integrated signal derived from an electronic stethoscope

Instantaneous heart rate was estimated using the median of S1-S1 intervals over 9 beats. The instantaneous heart rate HR_i in beats per minute at the ith beat is

$$HR_i = \frac{60}{f(T_{-4}, T_{-3}, \ldots T_3, T_4)}$$

where T_i is the S1-S1 interval between beat i and $i+1$ and f is the median operation.

For heart sounds recorded from the iPhone, in noisy segments where only either S1 or S2 was clearly visible, heart rate was calculated from whichever one was the most prominent heart sound in the record.

Results

The accuracy of heart rate extracted from audio data taken with the electronic stethoscope was compared to that from ECG by comparing QRS detection with S1 detection. All records examined had 100% R peak detection, totaling 997 beats. S1 detection was considered accurate if it was within 0.1 seconds of the corresponding R peak. Results are given in Table 1.

For audio data recorded with the iPhone, beat detection was verified by a human expert. Segments of audio data are shown in Figures 7 and 8 for clean and noisy signals respectively, which illustrate the feasibility of heart rate extraction from data recording using cellphones. However, the algorithm is robust to noise only when the inter-beat energy is smaller than that of S1 and/or S2 sounds. An example of this problem of dealing with ambient noise can be seen in the 4-5 second segment of Figure 9. Note also the respiratory amplitude modulation of the S1 peaks.

Positive predictivity	88.4%
Sensitivity	92.1%

Table 1: S1 detection results
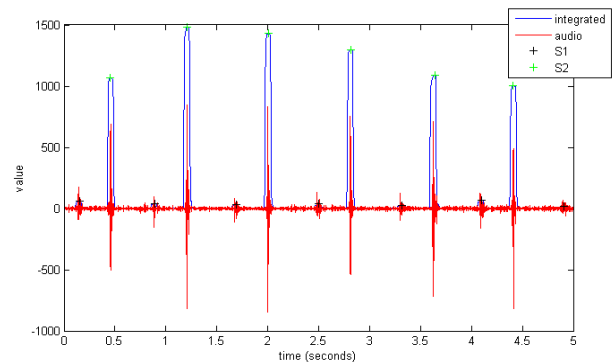
Figure 7: Heart sound detection from clean segment recorded from iPhone. HR=88 BPM

Figure 8: Heart rate detection from noisy segment recorded from iPhone with low ambient noise.

Figure 9: Heart rate detection from noisy segment recorded from iPhone with little high ambient noise.

Figure 10: Heart rate detection using HTC G1 with *ad hoc* stethoscope. Heart sounds are indicated by ▼

Figure 10 illustrates the performance of our detection algorithm on heart sound data recorded from the HTC G1. The first two heart sounds are labeled. Each following pair of peaks (▼) corresponds to pairs of S1 and S2 for each beat. Note that an *ad hoc* stethoscope was required to record these sounds, since the frequency response of the G1 HTC sound card below 100 Hz is so poor (see Figure 3). This stethoscope was constructed using a funnel with a plastic covering to mimic a conventional stethoscope.

A Telemedicine Infrastructure

Following rigorous validation of the technology on test patients, these methods will be integrated into Moca (mocamobile.org), a remote medical diagnostics platform aimed for use by rural healthcare workers in developing countries. The Moca framework currently consists of an Android client application that enables healthcare workers to upload rich media content of patient data from mobile phone devices to an OpenMRS electronic medical record backend system for review and diagnosis by an expert physician. By adding auscultation capabilities to the phone application, physicians will able to remotely listen to heart and lung sounds of the patients and more accurately diagnose patients with the appropriate heart and lung conditions. Over time, as more data is transmitted using Moca and stored in OpenMRS, a large collection of heart and lung sounds will accumulate and serve as a strong data set for the development of more sophisticated algorithms for automated detection of cardiac and pulmonary diseases on mobile phones in resource-poor regions.

Furthermore, by using the camera on the cellphone, it is possible to extend this infrastructure to create field databases to analyze, for example, skin lesions, eye diseases and wound infection (Celi *et al.* 2009).
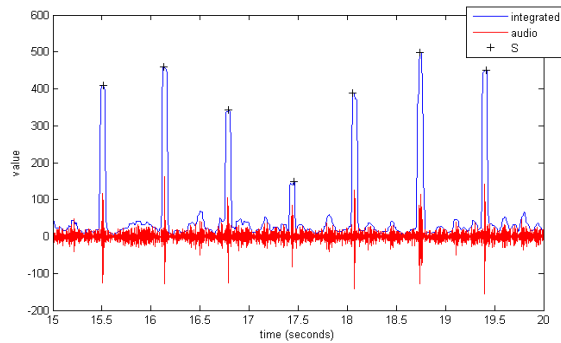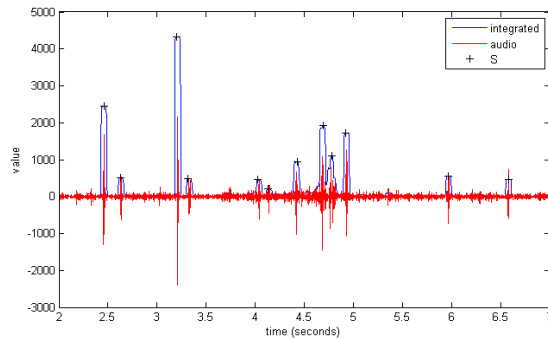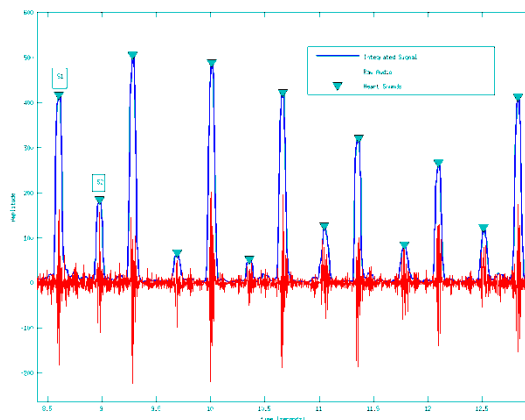
Discussion & Conclusions

We have developed a heart rate estimator (and heart sound locator) with a sensitivity of 92.1% and a positive predictivity of 88.4% for detecting each first heart sound using a gold standard database of ECG and heart sounds. Although this is well below the 99.9% levels reported for ECG beat detectors (Hamilton and Tompkins 1986), our algorithm is sufficient to detect the majority of the beats. If features of each heart sound need to be analyzed, then further signal quality checks on the morphologies and exact location (in time) of each heart sound will be required. Moreover, for heart rate estimation, the median approach means that the actual estimation is far more accurate than these figures represent.

When using cellphones to record heart sounds we found a high variance in quality between hardware, with some units being completely unable to record useful data because of the low frequency response characteristics of the sound card. Publicly available tests concerning sound card profiles indicated that the iPhone provides the best low frequency (<100 Hz) performance (which is key to cardiac auscultation), and our preliminary tests agreed with these conclusions. Preliminary tests on respiratory auscultation indicate that poorer performing sound cards are acceptable because information below 100 Hz is not essential.

Another important issue connected with cellphone auscultation is the problem of ambient noise and movement artifact. Figure 3 clearly illustrates this issue. Being able to identify artifacts and remove the affected

segments from physiological parameter estimation will be an essential part of any automated or semi-automated system such as described in Bhatikar *et al.* (2005). In Li *et al.* (2008), Li *et al.* (2009), and Nemati *et al.* (2009) we described a signal quality assessment approach for ECG, blood pressure and respiration respectively. In theory, it would be possible that this framework could be extended to incorporate audio auscultation signals and pulmonary signals. Future applications of the technique described in this paper could include detection of infections in fetuses (using HRV derived from recordings on the mother's abdomen), children and adults (using lung sound analysis). The key to success of these techniques is being able to inform the user when the recording location is providing sufficient signal quality to perform an accurate analysis of the data.

If the signal quality is too low, it may be possible to identify the underlying noise sources and remove them from the signal – i.e. filter the data, rather than remove noisy sections. However, since the noise overlaps in the time and frequency domain, filtering is extremely difficult. Mittra *et al.* (2009) used a second off-body microphone to record ambient noise to provide information for an adaptive filter. However, no comparable performance results are given for their work.

Another possibility for identifying and separating out the noise sources using a stereo microphone input, is to leverage independent component analysis (ICA). ICA removes statistically independent signal sources from the cardiac source if certain assumptions hold (Comon 1994). The effectiveness of ICA technique will depend on the accuracy of the assumption of linear, stationary *mixing* of the sources. We can see from Figures 6, 7, 8 and 10 that the S1 and S2 complexes exhibit slow changes in average energy over a period of several seconds. This observation is commensurate with the fact that as we breath, the location of the heart changes, possibly moving away and towards away from the microphone as we breath in an out (depending on microphone location). In these instances, the mixing matrix may no longer be stationary, and more complex de-mixing may be required, particularly for fetal heart sounds (Sameni *et al.* 2008).

The full value of capturing and analyzing cardiac and respiratory sounds is realized when it is integrated within a clinical information system. Probabilistic modeling to predict patient diagnosis and prognosis using physical (and even laboratory) findings almost always requires accompanying clinical history to optimize discrimination and calibration. To this end we have implemented a telemedicine framework (mocamobile.org) through which audio data can be uploaded and annotated, and expert evaluations and clinical treatment/follow-up recommendations can be rapidly sent to community health workers. Our next steps are to assemble such a database and make it publicly available.

In conclusion, we find that it is possible to re-task existing technology and hardware in resource poor environments to provide low-cost reliable diagnostic screening. We note that if a large medical device infrastructure and training system does become available in such locations, then other peripheral-based systems would form a complimentary diagnostic base to our proposed system. We also note that the diagnostic capability of the stand-alone cellphone is not confined to heart rate analysis, but can be extended to HRV, heart valve issues, lung function, infection, sleep structure and even depression (Sung *et al.* 2005) using acoustic and accelerometer inputs, for example. Moreover, fusing multiple independent signals to evaluate a given physiological function (such as cardiac activity), a more robust analysis of noisy field data can be made possible (Li *et al.* 2008).

References

Babybeat Inc., *Fetal Doppler for Monitoring at Home*. Web. 20 Oct. 2009. Online: http://www.babybeat.com/.

Bai, Y-W, and Lu, C-L., *Web-based remote digital stethoscope*. Proc. of IASTED International Conference, Internet and Multimedia Systems Applications, Grindelwald, Switzerland. 2005. 423-28.

Bhatikar, S.R., DeGroff, C., Mahajan, R.L., *A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics*, Artificial Intelligence in Medicine, March 2005, 33(3) 251-260, DOI: 10.1016/j.artmed.2004.07.008

Black, J., Sonenberg, L., and Scheepers, R. *Cell Phone Technology Is Key to Better Health in Africa,* Web. 20 Oct. 2009. http://research.microsoft.com/en-us/collaboration/focus/health/smartphone_clinical_diagnosis.aspx

Blad, S., Welin, A.-K., Kjellmer, I., Rosen, K.G., Mallard, C., *ECG and Heart Rate Variability Changes in Preterm and Near-Term Fetal Lamb Following LPS Exposure*, Reproductive Sciences 2008 15: 572-583.

Celi, L.A., Sarmenta, L., Rotberg, J., Marcelo, A., and Clifford, G.D. Mobile Care (Moca) for Remote Diagnosis and Screening. Journal Health Informatics in Developing Countries June 2009, 3(1):17-21.

Chen, J., Phua, K., Song, Y., and Shue, L. *Fetal Heart Signal Monitoring with Confidence Factor*, Multimedia and Expo, 2006 IEEE International Conference 1937-940.

Comon, P., *Independent Component Analysis: a new concept?*, Signal Processing, 1994, 36(3):287-314.

Dokur, Z. and Olmez, T., *Heart sound classification using wavelet transform and incremental self-organizing map*, Digital Signal Processing, Nov 2008, 18(6): 951-959.

Frasch, M.G., Muller, T., Weiss, C. Schwab, K., Schubert, H., and Schwab M., *Heart Rate Variability Analysis Allows Early Asphyxia Detection in Ovine Fetus*, Reproductive Sciences, May 1, 2009; 16(5): 509 - 517.

GSMArena Team. *Apple iPhone 3GS review: Same clothes, new feel*. Web. 20 Oct. 2009. http://www.gsmarena.com/apple_iphone_3gs-review-369p6.php

GSMArena Team. *T-Mobile G1 review: The whole cagoogle*, Web. 20 Oct. 2009. http://www.gsmarena.com/t_mobile_g1-review-337p7.php

Hamilton, P.S. and Tompkins, W.J., *Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database,* IEEE Trans. Biomed. Eng., BME-1986, 33(12):1157-1165.

Hutchinson, J. and Wyatt, J. *Freeplay Fetal Heart Rate Monitor*, Web. 20 Oct. 2009. http://www.designtoimprovelife.dk/index.php?option=com_content_custom&view=article&id=373:free-play-fetal-heart-reat-monitor&catid=9:winners-2009&Itemid=20

Jin, Z., Sun, Y., and Cheng, A.C., *Predicting Cardiovascular Disease from Real-Time Electrocardiographic Monitoring: An Adaptive Machine Learning Approach on a Cell Phone*, In Proceedings of the 31th International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBC), September 2009, Minneapolis, MN, in press.

Kovatchev, B.P., Farhy, L.S., Cao, H., Griffin, M.P., Lake D.E., Moorman, J.R. *Sample asymmetry analysis of heart rate characteristics with application to neonatal sepsis and systemic inflammatory response syndrome*. Pediatr Res. 2003;54 :892 –898.

Li, Q., Mark, R.G. and Clifford, G.D., *Artificial arterial blood pressure artifact models and an evaluation of a robust blood pressure and heart rate estimator.* BioMedical Engineering OnLine 8 July 2009, 8:13. doi:10.1186/1475-925X-8-13.

Li, Q., Mark, R.G. and Clifford, G.D. *Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter*, IOP Physiol. Meas. Dec 2008; 29: 15-32.

Mittra, A.K., *System simulation for a novel fetal monitoring methodology*, Int. J. Engineering Systems Modelling and Simulation, 2009, 1, 92-100.

Nemati, S., Abdala, O.T. and Clifford, G.D. *Data Fusion for Improved Respiration Rate Estimation*. MIT Technical Report, January 2009.

Oresko, J., Cheng, J., Duschl, H., and Cheng, A.C., *Transplanting a Resting ECG Machine to a Cell Phone for Real-Time ECG Acquisition, Feature Extraction, and Statistical Cardiac Summary Reports*, In Proceedings of the 2009 Biomedical Engineering Society Annual Scientific Meeting (BMES), October 2009, Pittsburgh, PA, in press.

Pan, J. and Tompkins, W.J, *A real-time QRS detection algorithm.* IEEE Trans. Biomed. Eng., BME-32 (3):230-236, 1985.

Richards, O. *UNICEF report: Half a million women die from pregnancy complications each year, World Socialist Web Site*. 20 Jan. 2009. Web. 30 Oct. 2009. http://www.wsws.org/articles/2009/jan2009/chil-j20.shtml

Rosqvist, T., E. Paajanen, K. Kallio, H. M. Rajala, and T. Katila. *Toolkit for lung sound analysis*. Medical and Biological Engineering and Computing 1995; 33.2:190-95.

R.A.L.E. Repository. Web. 14 Aug. 2009. http://www.rale.ca/.

Sameni, R., Jutten, C. , Shamsollahi, M.B, Multichannel *Electrocardiogram Decomposition using Periodic Component Analysis*, IEEE Trans. on Biomed. Eng., Aug. 2008, 55(8), 1935-40.

Sung, M., Marci, C. and Pentland, A., *Objective Physiological and Behavioral Measures for Tracking Depression*, Masters Thesis (excerpt), Massachusetts Institute of Technology, October 2005. http://vismod.media.mit.edu//tech-reports/TR-595.pdf

Syed, Z.H. *MIT Automated Auscultation System*, Masters Thesis. Massachusetts Institute of Technology, 2003.

Tan., A. and Masek, M., *Fetal Heart Rate and Activity Monitoring via Smart Phones*. mHealth Summit 2009, Washington DC, 29 Oct. 2009 – 30 Oct. 2009. Online: http://research.microsoft.com/en-us/events/mhealth2009/tan-masek-presentation.pdf

Thomas F. Budinger, *Biomonitoring with Wireless Communications*, Biomed. Eng. 2003. 5:383–412. doi: 10.1146/annurev.bioeng.5.040202.121653

United Nations. *It's Your World. United Nations Millennium Development Goals*, Web. 14 Aug. 2009. http://www.un.org/millenniumgoals/maternal.shtml

Universal Access- how mobile can bring communications to all. Intelecon for GSMA for Asia Congress, Singapore, Oct. 2006. Web. Oct. 2009. http://www.gsmworld.com/documents/universal_access_full_report.pdf