A Framework for Automated Heart and Lung Sound Analysis

Using a Mobile Telemedicine Platform

by

Katherine L. Kuan

Bachelor of Science in Electrical Engineering and Computer Science,

Massachusetts Institute of Technology (2009)

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

July 2010

Author.....
Department of Electrical Engineering and Computer Science
July 23, 2010

Certified by.....
Dr. Gari D. Clifford
Associate Director
Centre for Doctoral Training in Healthcare Innovation
University of Oxford
Thesis Supervisor

Certified by.....
Prof. P. Szolovits
Professor in Health Sciences and Technology and
Electrical Engineering and Computer Science
Thesis Co-Supervisor

Accepted by.....
Dr. Christopher J. Terman
Chairman, Department Committee on Graduate Theses

A Framework for Automated Heart and Lung Sound Analysis
Using a Mobile Telemedicine Platform
by
Katherine L. Kuan

Submitted to the
Department of Electrical Engineering and Computer Science

July 23, 2010

In Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

Many resource-poor communities across the globe lack access to quality healthcare, due to shortages in medical expertise and poor availability of medical diagnostic devices. In recent years, mobile phones have become increasingly complex and ubiquitous. These devices present a tremendous opportunity to provide low-cost diagnostics to under-served populations and to connect non-experts with experts. This thesis explores the capture of cardiac and respiratory sounds on a mobile phone for analysis, with the long-term aim of developing intelligent algorithms for the detection of heart and respiratory-related problems. Using standard labeled databases, existing and novel algorithms are developed to analyze cardiac and respiratory audio data. In order to assess the algorithms' performance under field conditions, a low-cost stethoscope attachment is constructed and data is collected using a mobile phone. Finally, a telemedicine infrastructure and work-flow is described, in which these algorithms can be deployed and trained in a large-scale deployment.

Thesis Supervisor: Gari D. Clifford
Title: Associate Director, Centre for Doctoral Training, Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford

Thesis Co-supervisor: Peter S. Szolovits
Title: Head of the Clinical Decision-Making Group, MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), Professor of Computer Science and Engineering, Department of Electrical Engineering and Computer Science (EECS), Professor of Health Sciences and Technology, Harvard/MIT Division of Health Sciences

Acknowledgments

First and foremost, I would like to express my gratitude to Brian and Tabetha Hinman and the MIT Public Service Center, who generously funded this work through a public service graduate fellowship. Without this funding and their continued encouragement, this thesis would not have been possible. Google also funded the work on the media viewer for OpenMRS and Sana through the Google Summer of Code (GSOC) 2009. Sana, which forms the backbone informatics system described in this thesis, has been funded by Telmex, the MIT PSC, Google, the Institute of Electrical and Electronics Engineers, the MIT International Science and Technology Initiatives, Vodaphone Foundation (through the Wireless Innovation Project), and the mHealth Alliance (formed by the United Nations, Rockefeller, and Vodafone Foundations).

In addition, I would like to extend a heartfelt thank you to the following people and organizations:

To Gari Clifford for mentoring me throughout the past one-and-a-half years (for GSOC, PSC Fellowship, class term projects, and my Master's thesis). I learned so much from him about the global health field and realities of healthcare technology adoption in resource-poor areas. He was incredibly helpful on my thesis, always available for questions, and taught me a great deal about the skills needed to become a successful scientific researcher.

To Peter Szolovits for his experienced advice and guidance on my thesis work, instruction during the HST Biomedical Computing course, and support of the Sana team.

To Tiffany Chen for her heart sound analysis code implementation, as well as time spent helping me capture heart sounds with various cell phones and preparing the paper for conference submission.

To Shamim Nemati for his code implementation and explanations on the autoregressive model method and signal purity algorithms (to be applied on lung sounds).

To John Guttag, Dorothy Curtis, Zeeshan Syed, and Gina Yi for providing me with access to the heart and lung sound data and files from previous research. To Gina for her code on lung sound analysis as well.

To Jose Gomez-Marquez for the opportunity to lead a lab activity in his Spring 2009 D-Lab Health class on building a stethoscope attachment for cell phones with his students.

To Hamish Fraser and Partners in Health for the experience of being a GSOC student for OpenMRS. That summer was my first glimpse at how a vibrant open source community can thrive with talented members who are open and willing to help each other out anytime.

To the MIT PSC (especially Sally Susnowitz and Alison Hynd) and the U.P. Manila National Telehealth Center (especially Dr. Alvin Marcleo) for the opportunity to do a fellowship in the Philippines during the Summer of 2009. It was undoubtedly one of the most thought-provoking and humbling experiences of my time at MIT.

To Sana for being the most inspiring team of developers, doctors, business people, faculty, and public health people. Being part of such a dynamic and passionate team has been such a wonderful experience. Thank you to Leo Anthony Celi, Richard Lu, and the other Sana doctors for always answering my questions about the medical context of this system and for providing feedback on the prototype as it evolved. To RJ Ryan for always being helpful and willing to pass on his incredible knowledge of software development to the other developers and me. To Jhonatan Rotberg for his help on the team, including UROP funding during my first semester on Sana and for providing hardware phone devices. To all the others on the team, especially the other developers, who aided me in many different ways.

To my dad, mom, and sister for their tremendous, unwavering support and love throughout this thesis, my MIT undergraduate years, and life as a whole. A big thank you all my family and friends who helped me from brainstorming for device instrumentation (especially my dad) to letting me record their heart and lung sounds.

I apologize if I inadvertently left anyone out, but so many people were helpful and I am grateful for their kindness.

Several open-source MATLAB toolboxes were used and adapted for this work: Gari Clifford's (2002) ECG toolbox, Tiffany Chen's (2010) heart sound segmentation algorithm, Gina Yi's lung sound analysis toolkit (2004), and Shamim Nemati's (2010) respiration analysis code. Descriptions and references to the original works are included in this thesis where they have been applied or modified.

The mobile phone frequency response curves in Chapter 4 (Figures 4-5 to 4-7 and Table 4-1) were included with permission from the GSMArena.com team with appropriate citations.

The percentile charts for evidence-based respiratory and heart rates in children compared to international ranges were also included with permission from Susannah Fleming (2010). These are located in Figures 1-2 and 1-5, as well as Tables 1-1 and 1-2.

The results on heart sound analysis were published earlier this year at the AAAI AI-D Spring Symposium in Stanford, CA from March 23-24, 2009. The article was called "Intelligent Heartsound Diagnostics on a Cellphone using a Hands-free Kit" with authors Tiffany Chen, Katherine Kuan, Leo Celi, and Gari Clifford (2010).

The project Sana was also published in the Journal of Health Informatics in Developing Countries in Volume 3, Number 1 in 2009 under Sana's previous name of Moca. The article was entitled "Mobile Care (Moca) for Remote Diagnosis and Screening" by Leo Anthony Celi, Luis Sarmenta, Jhonatan Rotberg, Alvin Marcleo, and Gari Clifford for the

Moca team. The conceptual design of the telemedicine infrastructure for capturing, labeling and transmitting resultant diagnostics back to the patient was developed by a large team of engineers at MIT from the Sana team. The system was designed by Drs. Celi and Clifford of MIT's Division of Health Sciences and Technology for the Edgerton Center class on Information and Communication Technologies for Developing Countries; ICT4D SP.716 (currently known as Nextlab) in the Spring of 2008 (MIT OpenCourseWare 2008). In the Fall of 2008, the client framework evolved from a nascent Symbian prototype developed by Andres Monroy-Hernandez to become an Android-based software system by the work of students RJ Ryan, Zack Anderson, and Boyuan Zhu in Professor Abelson's Mobile Programming Course. This system provided the ability to reliably transmit images, text, and voice recording to a back-end server in environments with intermittent connectivity. At the same time, a group of students in Nextlab (a class taught by instructor Jhonatan Rotberg) developed an OpenMRS-based backend to provide clinical review and the ability to send a diagnosis back to the uploading phone via SMS. Under the supervision of Dr. Clifford, I added several components to this system such as a rich media viewer in OpenMRS, upload of large medical ontologies into OpenMRS, upload and review of heart and lung sound data, and prototyped several security improvements for Sana.

Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Motivation

Many resource-poor regions of the world, especially in developing nations, are struggling to provide basic healthcare for people due to a number of reasons. There is often a shortage of trained medical specialists, poor access to diagnostic devices, and a supply-chain infrastructure that is unable to provide sufficient consumables, calibration, and maintenance for medical equipment (Malkin 2007). One example is in rural areas of the Philippines, where there is only one doctor for every 10,000 people, while the United States of America (U.S.A.) has one doctor for every 500 Americans (Nurse 2009). As a result, people who reside in such resource-poor communities are forced to travel long distances from their home to seek medical expertise in more urban areas. Apart from the significant medical bills, the travel is financially burdensome due to the transportation cost and lost earning potential.[1] Travel also includes the risk of a road traffic accident and is often uncomfortable to the point that it exacerbates a medical condition. In many cases, a decision to seek specialist care can be a life-altering decision that can utilize the majority of a family's savings and leave them in poverty (Whitehead et al. 2001).

The lack of proper medical attention for patients in developing nations has fatal consequences. In particular, acute respiratory infections (ARIs) are responsible for two

[1]Time that the patient and supporting family could have spent working.

million deaths per year and are the leading cause of mortality in children under five years of age in these countries (WHO 2009a). ARIs are caused by a number of factors including poor hygiene, overcrowding, low birth weight, malnutrition, and air pollution (especially from cooking) (Pandey et al. 1989). In 40 developing countries observed by UNICEF, less than half of the children with an ARI were actually taken to a recognized health provider[2] (UNICEF 2009). It is critical that when children first display the symptoms of an ARI, such as a cough or rapid breathing, that the appropriate diagnosis is made and treatment begins immediately. Otherwise, the risk of death increases dramatically.

Pneumonia is the most serious of the ARIs, with 5000 children dying from this condition per day, but it is fully treatable with antibiotics when the resources are available (UNICEF 2009). Without treatment, pneumonia can lead to high fevers, headaches, convulsions, or wheezing (UNICEF 2006). As fluid fills the lungs and breathing becomes difficult, the disease can be fatal for the individual. Unfortunately, the lack of trained community healthcare workers (CHWs) and lack of simple equipment mean that tens of millions of children remain undiagnosed, unable to receive appropriate treatment.

The burden of disease is not limited to ARIs in the developing world. As noted by the World Health Organization (WHO) in the recent decade, the major causes of deaths worldwide are shifting from infectious diseases to more chronic ones (Caldwell 2001). One group of diseases is called CRDs, or chronic respiratory diseases, which include asthma, obstructive sleep apnea, bronchiectasis, and pulmonary hypertension (WHO

[2] Where a recognized health care provider is defined as one found at a clinic or hospital or a community healthcare worker.

Africa 2009). Currently, hundreds of millions of people across the globe are believed to have CRDs (WHO 2010). In fact, a subset of CRDs called chronic obstructive pulmonary diseases[3] (COPD), are the third leading cause of death in the world with 3 million lives lost per year (WHO 2009c). There are 210 million people who have COPD today, and 90% of the cases specifically occur in low to middle income countries. Another pervasive CRD is asthma,[4] which affects an estimated 300 million people worldwide and causes 255,000 deaths per year (WHO 2008a). One of the symptoms of asthma is wheezing, which can be heard with a stethoscope and provide evidence to aid in diagnosis by the CHW (if he or she is properly trained). Timely diagnosis will help the patients be aware of the necessary steps to control the disease (i.e. avoid stimulating the asthma symptoms and obtain medication if possible) (WHO 2008a).

Even more alarming are the statistics for cardiovascular diseases (CVDs), which are the leading cause of death globally with 17 million deaths per year (WHO 2009b). This means that almost a third of all deaths in the world are due to CVDs, which include stroke, coronary, and rheumatic heart disease (Mackay et al. 2004). However, early detection of abnormal heart conditions can lead to behavioral and lifestyle changes to reduce the risk of heart attack[5] and stroke[6] through increased exercise, a healthier diet, and decreased tobacco use and exposure. Over 80% of the deaths from CVDs occur in low to middle income nations, which emphasizes the need for improved healthcare access for these people (WHO 2009b).

[3] Group of diseases that includes chronic bronchitis and emphysema. COPD is characterized by shortness of breath and chronic cough with sputum production.
[4] Occurs when the inner walls of the lung airways are inflamed by irritants (such as allergens in the environment). The airways become narrower, which limits air flow to the lungs to cause difficulty in breathing, coughs, or wheezing.
[5] Occur when blood is not able to reach the heart or brain because of blocked blood vessels (Mackay et al. 2004).
[6] When blood flow to the brain is interrupted (i.e. ruptured blood vessel in the brain) (Mackay et al. 2004).

Being able to ensure that patients are properly diagnosed is a difficult task. This is because training of healthcare workers and the supply of standard diagnostic medical devices (such as stethoscopes, ultrasound, or electrocardiographic machines) for health clinics are costly and logistically difficult. As mentioned before, resource-poor regions lack a supply chain infrastructure to support reliable healthcare equipment delivery. Even if a medical device is successfully delivered to a region, local adoption is still dependent on whether advanced training and expertise are required to operate the device, regional availability of replacement parts, and availability of technicians to service and calibrate the device.

Take for example the ARI timer that was distributed to CHWs in developing countries, in an attempt to reduce the number of deaths from pneumonia (Philips 2009). The ARI timer was designed to help determine whether a child exhibited rapid breathing, one of the first symptoms of pneumonia. The CHW counted the number of breaths that a child made in a minute, using the timer to determine when a minute had passed. Then the respiratory rate was compared against a look-up table to find the corresponding pneumonia classification level and determine the child's diagnosis. Even though the device was relatively inexpensive at $3.50 per unit, they only lasted for approximately a year. The battery was also hard to replace, so the entire timer was discarded for a new replacement device. As a result, ensuring a reliable supply of these ARI timers (in addition to devices for other diseases) to many remote regions is not a scalable or cost-effective solution.

Alternatively, there is a crucial need for a diagnostic tool that consists of equipment locally available in the country. Instead of a tool that can only perform a

single test, there is a need for a platform on which many types of diagnostic tests can be conducted for different heart and lung diseases. The core approach in this thesis is centered on the delivery of healthcare screening technology without requiring complex supply chain infrastructure and involving little additional training and cost from CHWs.

The rapid evolution of the wireless telecommunications industry can be leveraged in order to improve the capacity of CHWs to care for their local communities. In fact, mobile phones are experiencing rapid adoption across the globe with estimates that more than 80% of the world's population live within transmission and reception range of a cellular (mobile phone) communications tower (Fortner 2009). Furthermore, 64% of mobile phone subscribers are located in developing nations. Mobile phones are an obvious platform choice for several reasons other than their ubiquity. They can be used for telemedicine by sending patient data from remote CHWs over the cellular network to an expert for diagnosis. Mobile phones are also easy to keep charged (even if it is through the car battery), easy to conceal and protect from theft (because of their size), intuitive and familiar to use, and come replete with sensors such as cameras and microphones. The computing power of the mobile phone can also be leveraged so that the CHW can use the signal and image processing capabilities of the phone locally as a diagnostic tool (without any data transmission cost). In addition, mobile phones facilitate easy data back-up, security, authentication, tracking, and integration with medical records for auditing, longitudinal follow-ups, and appropriate allocation of resources.

The aim of this thesis is to develop a low-cost automated heart and lung sound diagnostic framework which uses a hands-free kit attached to a mobile phone, and a back-end centralized review system to allow expert annotation and rapid diagnostics. In

particular, calibrated methods (for automated or semi-automated signal processing) need to be developed to determine if such sound captured on mobile phones are of sufficient quality to accurately indicate infection or disease. The resulting low-cost system, as prototyped in the illustration of Figure 1-1, will hopefully address a growing and critical need in resource-poor settings for detection of heart and respiratory problems - particularly among fetuses (in the case of heart rate) and young children (in the case of respiratory sounds).

Ad hoc in-situ health diagnostics are likely to be appealing to those living in resource-poor locations because the modality of sourcing health care is already user-driven and *ad hoc* (Hall-Clifford and Chandler et al. 2010). For example, when a poorly educated and low-earning person feels the need to consult a health specialist, they often draw upon local vendors of medicines or social networks to source advice from those with little extra knowledge or training (Hall-Clifford 2005). In many cases, they will purchase inappropriate medications or health care solutions (Hall-Clifford 2005).

Figure 1-1. Schematic design of data capture and review system. A traditional stethoscope, used by an expert (left) is replaced by a simple hands-free kit attached to a cup (to amplify sounds) and a mobile phone to record and automatically process heart and lung sounds. Data are uploaded to a back-end server for quality auditing, and improving the classification algorithms.

Analyzing respiratory signals can help diagnose problems such as hyperventilation, apnea, asthma, lung infections, cardiac disease (such as tuberculosis pericarditis), or pulmonary edema. In addition, using the mobile phone as a diagnostic device has large cost-saving implications by equipping the most basic level of healthcare professionals with the facility to monitor patients' vital signs, perhaps track the progress of a fetus, and identify warning signs of adverse events (such as growth retardation through low heart rates). In addition to the benefits of avoiding travel, earlier detection of diseases would enable more rapid and successful treatments, reducing epidemics, and hence improving public health in general.

In order to develop a system that is specific to a given medical problem and geographical region, a large database of expert-annotated data samples is required. Therefore, a centralized remote data transfer, storage and annotation system is also necessary. To do this, an open source telemedicine client/server system known as *Sana*[7], (Clifford et al. 2008, Celi et al. 2009) was augmented to provide the appropriate infrastructure. Sana provides an Android-based Java application which allows authentication of users via barcode scanning, the entry or capture of categorical medical data together with any binary object (such as text, images, audio recordings, videos or files from peripheral systems connected by Bluetooth or USB), and an asynchronous upload method to a central database. An upload instigates a prompt to a medical professional to review the data and apply a medical code drawn from a relevant standard medical ontology, to categorize both the data and the follow-up course of action. (The system is described in Chapter 6.) Since multiple experts can review the data, accurate comparisons of expert diagnoses can be performed if a degenerate

[7] http://sanamobile.org/ - Previously known as 'Moca'

subset of alpha-numeric codes are used. The Android-based system for uploading and downloading data in the field can run on many different types of hardware, from laptops to mobile phones, and the data transfer can take place through any connection modality supported by the hardware (from 802.11x wireless to GPRS to Bluetooth).

The backend electronic medical record (EMR) system is based upon OpenMRS[8] an open-source EMR developed specifically for medical data collection in resource-poor areas, which allows for full transparency of the data, and provides an open application protocol interface (API) so that others can add additional functionality, such as data quality auditing or database integration.

One potential clinical scenario for the system described in this thesis is for the healthcare screening of women with young or unborn children. Access to cheaper or earlier detection of the diseases could lead to reductions in both child and maternal mortality, which are part of Millennium Development Goals 4 and 5 (United Nations 2009). An estimated 536,000 women died in 2005 during pregnancy or child birth, where 99% of those deaths occurred in resource-poor nations. In rural or peri-urban locations, semi-trained healthcare workers screen young mothers and pregnant women for childbirth complications and early signs of infection or disease. The healthcare workers' current method of determining fetal heart rate involves holding a cup to a pregnant woman's abdomen to listen and manually count the number of fetal heart beat sounds in a given period of time. This method is not dissimilar to pre-electronic times when a wooden "Pinard" stethoscope was used to perform the same task (Greer 2001). However, such an approach is prone to error due to background noise (both exogenous ambient and endogenous maternal/placental blood flow), as well as errors in timing,

[8] http://www.openmrs.org/

counting, or recording. The task therefore requires a moderate level of training and experience. Even for the trained human with very specific guidelines, inconsistency and inter-observer variability can be high (Neamatullah 2008). The employment of simple electronic devices by healthcare workers may provide significant gains in these areas (Blaya and Fraser 2006, Blaya 2008, Blaya et al. 2009).

The audio signal processing capability of the chipsets on mobile phones is designed to compress, filter, and transmit complex vocal signals. Audio recorded from the hands-free kit can therefore be automatically processed on the mobile phone to extract health metrics such as heart and respiratory rate. Automated or semi-automated algorithms may be able to provide the healthcare worker with decision support and training to more accurately detect the presence of infection or disease, and allow rapid referral to an appropriate treatment center. The medium-term goal of this project is to deliver and evaluate a prototype system with community partners in Central America and the Philippines in the future.

This thesis is organized into seven chapters. This first chapter provides motivation and background on current audio-based heart and lung sound analysis, together with an overview of existing mobile monitoring devices. Chapters two and three explore prototype algorithms for analyzing heart and lung sounds, respectively, on a mobile phone using standard databases. Chapter four provides an overview of device instrumentation required to analyze heart and lung sounds, plus an exploration of a prototype system for acquiring data in developing countries. Chapter five presents results of data collected using mobile phones, which by nature are a lower quality than the reference databases. Chapter six describes updates that were implemented to

enable a telemedicine system to review audio heart and lung data, as well as images and video (X-rays and ultrasound). Finally, chapter seven describes the main contributions in this thesis to the field of remote diagnostics, and possible future directions towards a full field trial.

1.2 Lung Sounds Background

1.2.1 Lung Anatomy and Clinical Background of Related Diseases

The lungs are the organs in the body responsible for respiration and are protected by the ribcage. Respiration is the process through which oxygen is supplied to the body and waste gases including CO_2 are expelled. When a person breathes, air enters through the nose or mouth and down the trachea (Moussavi 2006). The trachea, or "windpipe," then splits into two breathing tubes called the right and left bronchi. The bronchi, which lead to the right and left lung respectively, are further divided into smaller breathing tubes called bronchioles. The branching bronchioles end in tiny air sacs or alveoli, where there are an estimated 300 million alveoli in an average person's lungs (Scanlon and Sanders 2005).

The diaphragm is a muscle below the lungs separating it from the rest of the organs below. When the diaphragm contracts, the lungs expand and air is inhaled in a process called inspiration (Scanlon and Sanders 2005). Conversely, expiration occurs when the diaphragm relaxes, air leaves the lungs, and the lungs return to their relaxed position. The processes of inspiration and expiration repeat throughout the breathing

cycle. Respiration rate is defined as the number of breath cycles per minute (cpm), where a single cycle includes inspiration followed by expiration.

1.2.2 Respiration Rate

Traditionally, respiratory rate is measured by listening to lung sounds through a stethoscope, a device which will be described in more detail in Chapter two. Plethysmography is another method for measuring respiratory rate through observation of changes in chest volume (Mazeika and Swanson 2007), and pneumotachography measures the rate of airflow through the mouth (using a thermistor or air flow sensor) (Rao and Guha 2001). Respiration activity can also be derived from the electrocardiogram (ECG) and the photoplethysmograph (PPG) (Spyropoulos et al. 2008). Despite the multitude of options to monitor a patient's respiratory rate, many of these methods require advanced and relatively costly equipment. As a result, a *low-cost* method to capture lung sounds is needed.

Accurate respiratory ranges are crucial tools of reference in order for healthcare providers to identify abnormal individuals and diagnose them properly as having tachypnoea, for example, which is rapid breathing as an indicator of lung disease. According to one textbook, the respiratory rate ranges for normal healthy individuals are: 12 to 20 cpm for adults, 20 to 40 cpm for children, and 30 to 50 cpm for infants (Dreeben 2007, Jardins and Burton 2005). However, these values are rough estimates because there is no specific agreement in the literature on the acceptable ranges for respiratory rate among healthy people. Pre-defined ranges have been issued by international organizations such as the Pediatric Advanced Life Support (PALS), but the

scientific method that produced these numbers is unclear (Fleming et al. 2010, American Heart Association 2006).

Fleming et al. (2010) gathered the ranges suggested by various reputable health organizations and sought to identify evidence-based limits on respiratory rate (and heart rate) for newborns to children up to 18 years of age. They performed an extensive literature review from 1960 until 2009 to examine 69 studies from 20 countries, from which they were able to obtain respiratory rate data on 7,565 children and heart rate data on 150,080 children. Using a non-parametric kernel regression method on the data, they were able to produce respiratory rate percentile charts (in Figure 1-2) comparing results from actual data to the ranges provided by the Advanced Pediatric Life Support Group (APLS) (2004). The median respiratory rate (based on the actual data) is shown in green in Figure 1-2. The 99^{th} percentile curves in blue often lie far outside the proposed international ranges in red. These results imply that organizations such as the APLS have too coarse an estimate on the valid ranges of respiratory rate in children from infancy to 18 years old. Based on studies in the literature, the standardized ranges should actually be expressed as nonlinear curves with the limits as provided in Table 1-1 by Fleming et al. (2010).

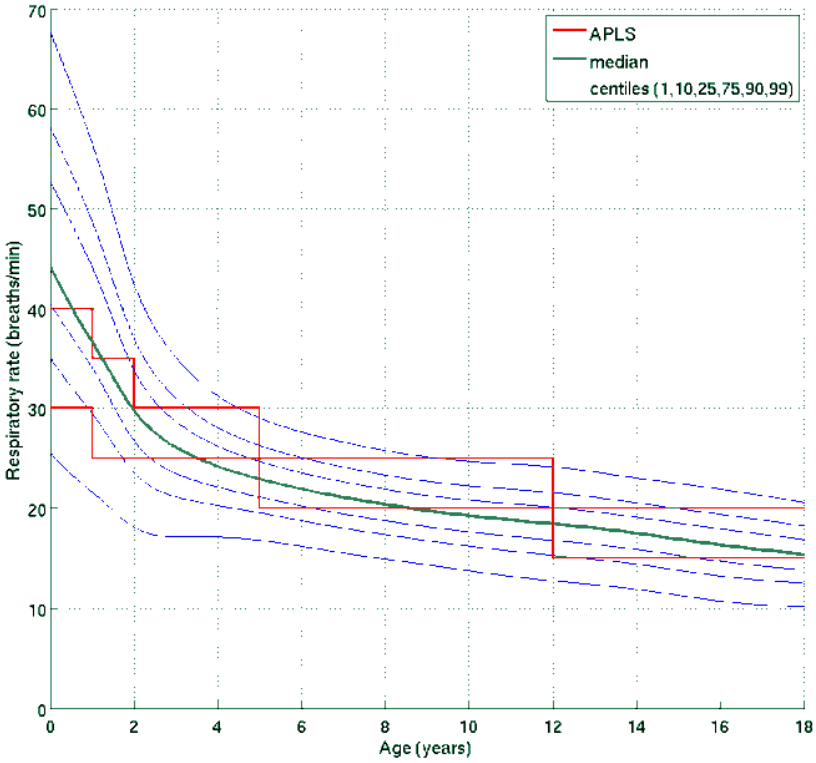


Figure 1-2. Percentile charts for evidence-based respiratory rates (blue) compared to the international ranges (red) from APLS. Reproduced from Fleming et al. (2010) with permission.

Age Range	1st centile (cpm)	10th centile (cpm)	25th centile (cpm)	Median (cpm)	75th centile (cpm)	90th centile (cpm)	99th centile (cpm)
0 – 3m	25	34	40	43	52	57	66
3 – 6m	24	33	38	41	49	55	64
6 – 9m	23	31	36	39	47	52	61
9 – 12m	22	30	35	37	45	50	58
12 – 18m	21	28	32	35	42	46	53
18 – 24m	19	25	29	31	36	40	46
2 – 3y	18	22	25	28	31	34	38
3 – 4y	17	21	23	25	27	29	33
4 – 6y	17	20	21	23	25	27	29
6 – 8y	16	18	20	21	23	24	27
8 – 12y	14	16	18	19	21	22	25
12 – 15y	12	15	16	18	19	21	23
15 – 18y	11	13	15	16	18	19	22

Table 1-1. Evidence-based respiratory rate ranges for healthy children from 0 to 18 years old. Reproduced from Fleming et al. (June 2010) with permission.

1.2.3 Classes of Respiratory Sounds

This section describes respiratory sounds as a field of interest because they serve as important indicators of respiratory-related diseases (Pasterkamp et al. 1997). Respiratory sounds can be split into two categories: normal and adventitious[9] sounds. In general, lung sounds occur at low frequencies from 100 Hz to 1200 Hz (while the human audible frequency range is from 20 Hz to 20,000 Hz) (Pasterkamp 2008). Normal lung sounds usually occur below 500 Hz, while adventitious lung sounds are heard above 500 Hz. Researchers often apply high-pass filters, with a filter cutoff frequency around 60 Hz, to lung sound recordings to remove noise and improve the signal clarity (Hadjileontiadis 2009).

Normal breath sounds can be categorized into four classes: tracheal, bronchial, bronchovesicular, and vesicular sounds (White 2006). Each class of sounds is detected during auscultation according to the characteristics described below:

- **Tracheal**: Heard over the trachea as harsh, high-pitched, and discontinuous sounds (McCann et al. 2005b, Pasterkamp et al. 1997). Originate from turbulent air flow in the upper airways of the body and cover a wide frequency range from less than 100 Hz to 1500 Hz (Fenton et al. 1985, Gavriely et al. 1981). Expiration phase lasts slightly longer than inspiration (White 2003).
- **Bronchial**: Normally heard over the manubrium or upper part of the sternum (Cohen and Landsberg 1984). Consist of high pitch and high amplitude sounds similar to the sound of air blowing through a tube. There is a brief pause between

[9] Abnormal sounds indicating the presence of a pulmonary disease

expiration and inspiration, where expiration is three times as long as inspiration (White 2003, McCann et al. 2005b). Abnormal if heard over the entire lung (where vesicular sounds should be heard), indicating that the lung is consolidated (Landois 1891).

- **Bronchovesicular**: Heard over the anterior chest between the scapulae and sternum, between the first and second intercostal spaces (White 2003). They possess pitch and intensity characteristics midway between vesicular and bronchial sounds. They can be heard during the inspiratory and expiratory phases, which each last for approximately the same duration and do not have a pause in between them (McCann et al. 2008, White 2003).
- **Vesicular**: Soft and low-pitch breezy sound heard over the majority (and particularly over the base) of the lung (McCann et al. 2005b). Occurs as a result of changing airflow patterns in the lungs. Inspiration is more audible than expiration (White 2004). There is no pause between the two phases, but inspiration is three times as long as expiration (White 2003). Similar to the sound of wind blowing through a tunnel.

Tracheal sounds are louder than normal (vesicular) lung sounds in that both inspiration and expiration phases are audible (Sovijärvi et al. 2000). Tracheal sounds are also directly correlated with air flow, so respiratory rate can be extracted in a straightforward manner. However, these sounds are typically not used for diagnostic purposes because respiratory disease symptoms may not be prevalent in the upper areas of the chest. Instead, for diagnosis, physicians focus on lung sounds

(bronchovesicular and vesicular sounds) heard on the anterior and posterior chest, which unfortunately are usually less audible and at lower frequencies compared to tracheal sounds with a peak around 100 Hz (Cohen and Landsberg 1984, Pasterkamp et al. 1997). Ausculatory signals can be analyzed to extract respiratory-related information such as breathing rate, phase onsets, and disease-specific features which provide evidence for clinical diagnosis decisions (Moussavi 2006, Pasterkamp et al. 1997, Yi 2004).

Adventitious lung sounds signify a pulmonary disorder in patients and are categorized as continuous (i.e. wheezes) or discontinuous sounds (i.e. crackles). The distinct natures of each class of lung sounds show promise that sophisticated signal processing algorithms will be able to provide automatic classification of the sounds. Qualitative details on the differences between adventitious sounds are provided below.

- **Crackles (Rales)**: Discontinuous and explosive popping sound usually during inspiration (McCann et al. 2002). Caused by air passing through moist airways and alveoli. Short duration crackles are called fine crackles, which occur at higher frequencies with a high pitch and low amplitude (Murray and Nadel 1995, Bates 1995). Fine crackles are usually heard during inspiration (5-10 ms duration) and can indicate COPD, pneumonia, pulmonary fibrosis,[10] or an early sign of heart failure (Yeginer and Kahya 2005, White 2004). Long duration crackles (20-30 ms long) are called coarse crackles. They are usually low-pitched and associated with pneumonia, bronchitis,[11] atelectasis,[12] or pulmonary

[10] Scarring of the lungs resulting in thickened lung tissue (MedlinePlus 2009).
[11] Inflammation of the bronchi primarily caused by cigarette smoking. Also accompanied by a cough with sputum production (Cabot 1912).

edema.[13] With respect to timing, crackles can occur during the start of inspiration (chronic bronchitis and emphysema), midway through inspiration (bronchiectasis),[14] or end of inspiration (pulmonary fibrosis) (Souhami and Moxham 2002).

- **Pleural Rub**: Produces a sound similar to two pieces of leather rubbing together and caused by inflammation of the pleural membrane. Pleural rub can be a symptom for pleural effusion,[15] pneumonia, lung abscess,[16] or tuberculosis[17] (White 2004).
- **Rhonchi**: Continuous and low-pitched snoring sounds similar to wheezes. Usually at frequencies less than 300Hz with a duration longer than 100 ms (Davis et al. 1999). Heard primarily during expiration over central airways of the body when airways are filled with fluid or mucus secretions (McCann et al. 2008).
- **Squawks**: Short wheezes on inspiration that last from 50 to 400 ms (Pasterkamp et al. 1997). Has musical quality to the sound. Occurs frequently in patients with pneumonia, pulmonary fibrosis, or interstitial[18] lung diseases (Earis 1992).

---

[12] Collapsed alveoli in lungs (Hoffmann et al. 1902).
[13] Excess fluid in the lung tissue that can cause cough, crackles, wheezes, rapid breathing, or irregular heart beat (White 2004).
[14] When the bronchi are dilated, susceptible to infection, and cause reduced airflow (White 2004).
[15] Characterized by excess fluid in the pleural membrane between the lungs and the chest wall (Smeltzer et al. 2009).
[16] Build up of pus in the lungs (WHO).
[17] Infection of the lungs by Mycobacterium tuberculosis that causes alveoli inflammation (White 2004).
[18] A group of lung disorders involving scarring (or fibrosis) of the tissue around the air sacs and inflammation of the air sac walls (University of Maryland 2006).

- **Stridor**: Harsh crow-like sounds that occur during inspiration from obstructed upper airways due to conditions such as a tumor or foreign body obstruction (Ashby 1905, White 2004).
- **Wheezes**: High-pitched whistling or musical sounds caused by narrowed lung airways (McCann et al. 2002). They create a few sharp peaks in the power spectrum of the lung sound, with a dominant frequency around 400 Hz (but could be in the range of 100 Hz to 1200 Hz) (Qiu 2005, Fenton et al. 1985). A wheeze has duration of longer than 100 ms and is usually heard in expiration. These adventitious sounds indicate diseases such as pneumonia, asthma, chronic bronchitis, emphysema,[19] congestive heart failure, or bronchiectasis (Qiu 2005, Murray and Nadel 1995, Bates 1995). A monophonic wheeze contains a single frequency, while a polyphonic wheeze contains multiple frequencies. The latter can be heard during the expiration phase of patients with asthma, emphysema, or chronic bronchitis (Souhami and Moxham 2002).

The challenge in detecting abnormal lung sounds is that they do not occur in isolation. Often, respiratory diseases will involve multiple types of these abnormal lung sounds. There is also a limited availability of labeled lung sounds data from unhealthy patients with respiratory diseases. Sana will therefore be used to capture lung sounds recordings from healthy and unhealthy patients observed in real field situations by healthcare workers. The resulting collection of lung sound recordings will enable the future development of signal processing algorithms for automated disease detection.

[19] Shortness of breath primarily caused by cigarette smoking or air pollution (Lewis 1999).

Before this can be done, it is necessary to design the system such that high enough quality data can be captured.

1.2.4 Existing Research in Automated Lung Sound Classification

Expert use of a stethoscope to perform auscultation[20] has been a longstanding method for diagnosing patients with respiratory diseases. However, computerized analysis of lung sounds offers new advantages: the ability to maintain a record of the patient's lung sounds, achieve objective interpretation of the data (instead of a subjective one from a doctor), and to extract information from a signal that may be outside the hearing range of humans (possibly missed during expert auscultation) (Qiu et al. 2005).

Gavriely et al. (1994) found that lung sound analysis revealed additional useful information needed to diagnose a patient with a respiratory disease. The study involved 493 respiratory health screening patients (385 males and 108 females from age 21 to 77), where 91 had pulmonary diseases including bronchitis and obstructive lung disease. Gavriely et al. tested the effectiveness of adding a computerized lung sound analysis test in conjunction with an existing pulmonary function test screening method called spirometry.[21] Lung sounds were captured with piezoelectric sensors from four parts of the body: the trachea, right anterior chest, lower left posterior, and right lower posterior chest. Of the 91 patients with actual respiratory diseases, 74 had abnormal spirometric results and 54 patients were found to have abnormal lung sounds. With a

[20] The act of listening to the body's sounds (lungs, heart, trachea, gastro-intestinal tract) for disease detection (Anders 1907, Anders and Boston 1911).
[21] Spirometry measures maximum air volume or airway resistance in the lungs (Gavriely et al. 1994, Cohen 1990)

combination of both screening methods, Gavriely et al. demonstrated an increase in pulmonary disease detection from 71% sensitivity (with spirometry alone) to 87% (with both methods). In some cases, the computer analysis tool was able to detect abnormal lung sounds (for example, a wheeze that only occurred once or twice in 15 minutes) that were overlooked by three experts listening to the same recordings.

Wheezes are abnormal breath sounds that are periodic, with a dominant frequency greater than 100 Hz, lasting for longer than 100 ms (Qiu et al. 2005). Beck et al. (1992) detected wheezes by interpreting a peak in the frequency domain that was greater than 15 times the mean signal power. Other researchers developed empirical rules and scoring systems to define a wheeze as a certain peak in the frequency domain (Fiz et al. 2002). Qiu (2005) developed an algorithm where the threshold for wheeze detection was frequency and duration-dependent using an auditory model. In his algorithm, Qiu identified a peak in the frequency domain and compared it against a threshold to verify that it was a single peak. The threshold value was frequency dependent and calculated from the energy in the auditory filter bandwidth (instead of the total power). Tracheal breath sounds from 36 patients were obtained for testing, where 11 were verified to exhibit wheezing according to expert consultation. The algorithm was able to detect wheezes in 10 of those patients, but more detailed and quantifiable performance results were not provided in the article.

Crackles are another class of abnormal lung sounds of interest to researchers. In the cases of interstitial lung diseases and heart failure, crackles are first heard at the base of the lungs and move upward as the disease worsens. The pitch, spectrum, and timing of the crackles are different in each type of disease, which show promise for

automated classification methods. In addition, counting the number of crackles can indicate the severity of a disease. Murphy et al. (1989) developed an automatic counter of crackles for the computer and tested the algorithm on 100 lung sound recordings from 41 patients. The automated crackle estimator was strongly correlated ($R = 0.74$ and $p < 0.001$) with the values obtained by physicians who manually counted the crackles. Kaisla (1991) developed a different crackle counter based on peak detection in a lung sound sonogram.[22] The method was applied to recordings that contained 117 breath cycles with 1064 crackles from 20 unhealthy patients - 10 with fibrosing alveolitis (FA) and 10 with bronchiectasis (BE). The counter's results were compared to the number of crackles identified by two observers looking at time-expanded waveforms of the recordings. The algorithm yielded a sensitivity of 89% for patients with FA and sensitivity of 80% for patients with BE. For more advanced methods, Yeginer and Kahya (2005) used a wavelet network to characterize crackles within a lung sound signal by optimizing two weight factors, scaling, time-shifting, and frequency parameters. The crackles were identified as fine or coarse crackles based on likelihood probabilities from Bayes' rule. On a database of 238 coarse crackles and 153 fine crackles from two patients (as labeled by medical experts), the algorithm was able to correctly classify coarse crackles with 69% accuracy and 84% for fine crackles. One of the reasons why a distinction was possible was because fine and coarse crackles are observed to occur at different times within the inspiration process (Moussavi 2006). Because crackles have a short duration and low intensity that is difficult for a human to hear in traditional auscultation, computerized analysis of crackles can provide useful information in the early detection of lung pathologies (Kaisla 1991).

[22] Signal frequency distribution over time (Kaisla 1991).

Even simple lung sound analysis such as respiration rate extraction provides valuable information about the patient's condition. Cretikos et al. (2008) described respiratory rate as "the neglected vital sign." Hoping to motivate doctors and nurses to change the current poor documentation practices of respiration rate in hospitals, Cretikos emphasized the significance of abnormal respiratory rate as a precursor to major clinical events such as cardiac arrest or admission to the Intensive Care Unit. Adult hospital in-patients were classified as abnormal for a respiration rate higher than 20 cpm, critical condition for rates higher than 24 cpm, while a rate higher than 27 cpm was one of the most important signs preceding cardiac arrest. Consequently, building a system that performs automated respiratory rate extraction can be very valuable.

Detection algorithms such as those described above are moving into the clinical setting. For example, a new automated and portable cough analyzer (CA) was created by Krajnik et al. (2010). The CA was tested by 13 patients with a chronic cough (8 females and 4 males from 19 to 58 years of age) by wearing the monitoring device (with an audio recorder) around their chest for 24 hours. The CA registered information on the number of cough incidents and time spent coughing per hour on flash memory, which was later transferred to a computer for analysis. The counts from the CA were compared against subjective evaluations from the patients using a cough severity scale from 1-10. Only a moderate correlation was found between subjective cough score and the number of cough incidents per hour with $r = 0.60$ and $p = 0.029$ from the Wilcoxon[23] rank test. The CA was also able to provide information about the patients' conditions beyond what could be observed during a doctor visit. For example, Krajnik et al. found

[23] To test a hypothesis against the median of a dataset when the population may not be normally distributed (Utts and Heckard 2005).

that more cough incidents seemed to occur during the day with a median 45.5 cough incidents per hour during the day and 8 at night. They also found that more time was spent coughing during the day with a median 20 seconds per hour spent coughing during the day versus 3 seconds per hour at night. Ultimately, this work allows doctors to gain objective information on the actual severity of the cough based on the 24-hour monitoring data (instead of the patient's subjective opinion) in order to more accurately diagnose the patient. However, the sample size of the study was small and three patients did not finish the study, while eight patients removed the device early due to discomfort or interference with job responsibilities. Hence, the prototype can still be evolved further to ensure full patient compliance.

Another example of a clinical use case for automated signal processing methods is a toolkit for non-invasive lung sound analysis developed by Rosqvist et al. (1995). A repository of 30 crackling and 10 wheezing lung sounds was created from patients with fibrosing alveolitis, bronchiectasis, chronic obstructive pulmonary disease, and heart failure. Rosqvist et al. used two condenser microphones placed on the torso at the base of each lung for recording, and the microphones were enclosed in a protector cup with rubber straps to reduce ambient noise. A high pass filter with a cutoff frequency of 50 Hz was used to filter out low frequency noise. Their detection algorithm was demonstrated to have a classification accuracy of 88% average positive predictivity (PPV) on patients with fibrosing alveolitis and 83% PPV for patients with bronchiectasis. A different software suite for acoustic respiratory analysis was also developed in MATLAB by Yi and Guttag (2004). Their software characterizes lung sounds using a respiration rate detector, respiratory phase onset detector, respiratory phase classifier,

along with a crackle and wheeze detector. The tracheal breath sound respiratory rate algorithm from Yi's work is described in more detail in Section 3.2.1.

Regardless of the type of disease or lung sound for which an algorithm is developed, the signal being analyzed must be recorded with sufficient signal-to-noise ratio to extract useful information. Thus, it is important that even during noisy conditions in health clinics, the recording of the signal is still clear, while additional signal processing can eliminate the remaining external noise. One way Rosqvist et al. (1995) addressed this issue was to put another microphone near the patient and record the ambient noise as a reference signal for filtering later. The noise was detected to be around less than 50 Hz, so high pass filtering was a good option to improve signal quality. The use of stereo microphones would be another appropriate way to deal with noise.

Another study on noise reduction was done by Chang and Cheng (2008), who explored which of three methods (AR model, Mel-frequency cepstrum, and bispectrum diagonal slices) would perform the best in lung sound recognition with the presence of varied degrees of noise. They obtained four original 20-second recordings of normal tracheal sounds, wheezes, and crackles sampled at 8 kHz from the Stethographics Inc. (2007) website. They added various types of noise (white, babble[24], and car noise from the NOISEX database) to the lung sound recordings to produce test signals with an SNR of 0, 5, 10, 15, 20, 30, and 40 dB (Varga et al. 1992). Each test signal was divided into 512-sample segments to determine the corresponding bispectrum diagonal slices[25]

[24] Speech babble noise occurs when many speakers in the background contribute to the noise in the signal (Hansler and Schmidt 2008).
[25] Involves taking diagonal slices ($w_1 = w_2$ in the w_1 x w_2 plane) of the bispectrum of the signal, which is the 2D Fourier transform of the third order cumulant of the signal (Chang and Cheng 2008).

(BDS), 13th order Mel-frequency cepstral coefficients[26] (MFCC), and 8th order autoregressive (AR) model coefficients for that segment. Dynamic time warping was used to measure time series similarity between the signals defined by the three feature sets and the original lung sound signal without noise (Ye 2004). The BDS method ended up having a better mean recognition rate for lung sounds than the AR method by 20% and MFCC method by 31.4%. Meanwhile, the MFCC method was shown to be very susceptible to noise because its recognition rate dropped the fastest out of the three methods as the SNR level dropped. Similarly, Suzuki et al. (1995) were interested in extracting useful lung sounds recorded in noisy environments. They developed a real-time adaptive noise canceller by integrating the system with a digital signal processor and custom electronic stethoscope. By using an adaptive filtering technique to cancel noise on the recordings, they were able to demonstrate a 30 dB reduction in noise approximately within five seconds.

While the previously described studies attempted to reduce noise from the external environment, Yip and Zhang (2001) addressed the problem of interference in lung sound recordings from bodily noises such as heart sounds. They created a platform that consisted of an electronic stethoscope and adaptive algorithm with automatic gain control. They successfully reduced the presence of heart sounds in lung sound recordings by values within the range of 75% to 83% depending on the specific recording location on the chest.

Only one study has reported the use of mobile phones to record lung sounds. Anderson et al. (2001) examined whether breath sounds could be recorded with a

[26] Used to describe a signal's frequency characteristics where the coefficients are calculated based on the discrete cosine transform of the logarithm of the signal's power spectrum (Diederich 2008, Chang and Cheng 2008).

mobile phone to determine whether patients had asthma or not. Tracheal breath sounds were collected from 20 patients from 12 to 61 years old with an equal number of females and males. Based on clinical history, peak flow recordings, and lung function, seven patients had asthma and were being treated, while the rest were considered to be in normal condition. The microphone of the phone was placed directly over the trachea for capturing five breath cycles for each patient. The recordings were captured through a voice call to an eReceptionist automated voicemail service,[27] which sent the recording in GSM format 6.10 (with 16-bit resolution and sampling frequency 8 kHz) to a computer for analysis. The type of phone and soundcard on the phone were not specified, so it is unclear whether the signal was distorted by the phone during the recording process, by transmission across the cellular voice network, or by the voicemail recording service's data compression methods. While Anderson et al. concluded that recording breath sounds on a mobile phone and processing them on a computer was effective enough to distinguish asthma patients from healthy ones, there are several limitations to the study. The results were only quantified to the extent that abnormalities in the spectrogram were correlated with asthma using Fisher's exact test[28] (with $p < .0001$). In addition, they only focused on one respiratory condition: asthma, and for very few subjects. It is unlikely that the technique would work in the real world because ambient noise would play a significant role in error rates, the cellular voice network would distort the recordings, and patients are likely to have infections and manifest many different sounds to the small data set they analyzed. This thesis addresses these problems, as a prerequisite for developing automated analysis

[27] http://www.yac.com
[28] To test if two treatment groups are classified according to two variables of interest in different proportions (Osborn 2000).

techniques for a range of respiratory diseases, over different recording systems and in different environments.

There are also several other recent projects which attempt to use the mobile phone as a general diagnostic platform for healthcare workers in developing countries. Abeyratne et al. explored whether low-cost electronic devices, including mp3 players and mobile phones, could be used to record cough and sleeping sounds to help in the diagnosis of pneumonia and sleep-related diseases (Bill and Melinda Gates Foundation 2010). No results have been reported yet on the diagnostic accuracy of their iPhone application. Black et al. (2009) focused on developing mobile phone applications for use by semi-trained health care workers in Mozambique, Africa. As part of the Nossal Institute of Global Health, they focused on creating applications that do not incur recurring costs by avoiding calls, SMS, and data transfer on the phones. Although this makes sense in financial terms, the ability to create expert-labeled databases from field data is far more difficult in such a situation. They developed an array of applications including a gestational dates' calendar for pregnant women, drug dose calculator, drip rate calculator, and drug reminder alarm. They also have simple respiratory rate and heart rate calculators which operate similar to a stopwatch as the healthcare worker manually counts breath cycles or heart beats for a certain period of time. The application does not take patient sound or monitoring signals as input, so the system relies on the healthcare worker being able to accurately count heart beats and breath cycles manually - an assumption that does not always hold true. Moreover, the system is susceptible to the vagaries of human estimation and transcription.

From this literature review, it can be observed that there are many problem-specific solutions which are not easily generalizable. For example, there are sophisticated digital techniques for extracting clinically relevant information from lung sounds, but they depend on a microphone, digitizer, and specific desktop application. Conversely, there are solutions designed for rural healthcare workers for simple functions such as respiratory rate counting, but do not aid in respiratory disease diagnosis. The mobile phone platform holds much potential for use as a diagnostic tool that is simple, non-invasive, relatively inexpensive, and able to output objective and quantitative analysis on the respiratory health of patients. The approach outlined in this thesis is to create an integrated solution to all of these challenges by delivering a robust mobile phone platform, which can record and provide analysis of the lung sounds, with little training required by the caregiver.

1.3 Heart Sounds Background

Closely related to the issue of lung sound analysis is that of heart sound analysis. As pointed out above, heart and lung sounds are actually sources of noise for the other's analysis (Yip and Zhang 2001). However, heart sounds are generally of lower frequency and therefore the analysis of each signal source separately is often possible (Piinla 1995). Section 1.3 now reviews the field of heart sound analysis.

1.3.1 Heart Anatomy

The heart is responsible for pumping blood to the organs and cells throughout the body to provide them with the nutrients and oxygen needed for survival (Taylor 2000). The heart is comprised of four chambers. The upper left and upper right chambers are called atria, while the lower left and lower right chambers are called ventricles. The chambers are separated by a muscle called the septum, and there are multiple valves to connect the various chambers to each other. Oxygen-deficient blood from the right atria flows to the right ventricle through the tricuspid valve and then gets pumped through the pulmonary valve to the lungs in order to gain oxygen (Sherwood 2009). The oxygen-rich blood returns to the heart into the left atrium and gets pumped through the mitral valve to the left ventricle. Finally, the blood flows throughout the rest of the body when the aortic valve opens, and the left ventricle compresses to pump the blood. Blood travels to the organs and cells of the body by means of a network of arteries, arterioles, and capillaries, while the oxygen-deficient blood travels back to the heart through a network of veins and venules. Altogether, the heart and circulatory system of arteries and veins make up the body's cardiovascular system.

The heart beat is a two-step process. The "diastole" step begins when the atria fill with blood, the sinoatrial (SA) node emits an electrical signal, and the tricuspid and mitral valves open to allow blood flow into the ventricles (Taylor 2000). During diastole, the ventricles relax (McCann et al. 2005a). The second step, which is called the "systole," occurs when the ventricles fill up and the tricuspid and mitral valves start to close. Then ventricles start to contract, and both the pulmonary and aortic valves open to allow blood flow to the lungs and rest of the body.

1.3.2 Heart Rate

The heart rate, also known as pulse rate, can then be defined as the number of heart beats per minute (bpm). The pulse can be measured manually by placing two flat fingers (not including the thumb) on the wrist, neck, or temple, where main arteries are located (King and Reiss 2001). According to one textbook, the range of acceptable heart rates for an average person at rest is claimed to be: 60 to 100 bpm for an adult, 50 to 100 bpm for a teenager, 74 to 140 bpm for a child, and 70 to 170 bpm for an infant (Dreeben 2007). As introduced in Section 1.2.2, Fleming et al. (2010) obtained heart rate information on 150,080 healthy children and produced the chart pictured in Figure 1-3 comparing the percentiles of the actual data (blue) with the international ranges (red) of APLS. Again, the 1st and 99th percentile heart rate curves were not sufficiently described by the step-wise linear ranges. These results implied that children may have not been diagnosed properly for conditions such as tachycardia (abnormally fast heart beat) if the upper ranges of heart rate for normal healthy children are incorrectly defined by APLS and other organizations (Fleming et al.). Based on this research, Fleming et al. proposed new acceptable heart rate ranges as provided in Table 1-2.

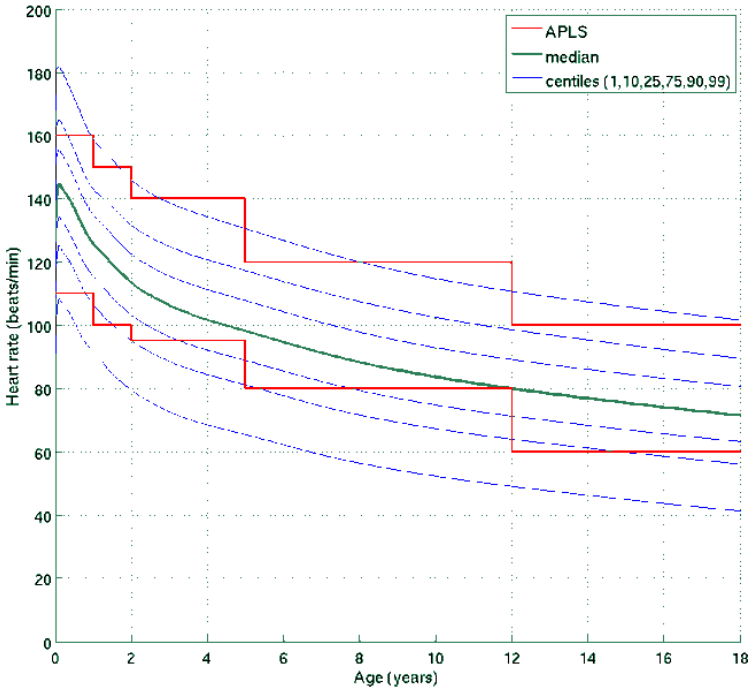


Figure 1-3. Percentile charts for evidence-based heart rates (blue) compared to the international ranges (red) from APLS. Reproduced from Fleming et al. (2010) with permission.

Age Range	1st centile (bpm)	10th centile (bpm)	25th centile (bpm)	Median (bpm)	75th centile (bpm)	90th centile (bpm)	99th centile (bpm)
Birth	90	107	116	127	138	148	164
0 – 3m	107	123	133	143	154	164	181
3 – 6m	104	120	129	140	150	159	175
6 – 9m	98	114	123	134	143	152	168
9 – 12m	93	109	118	128	137	145	161
12 – 18m	88	103	112	123	132	140	156
18 – 24m	82	98	106	116	126	135	149
2 – 3y	76	92	100	110	119	128	142
3 – 4y	70	86	94	104	113	123	136
4 – 6y	65	81	89	98	108	117	131
6 – 8y	59	74	82	91	101	111	123
8 – 12y	52	67	75	84	93	103	115
12 – 15y	47	62	69	78	87	96	108
15 – 18y	43	58	65	73	83	92	104

Table 1-2. Evidence-based heart rate ranges for healthy children from 0 to 18 years old. Reproduced from Fleming et al. (June 2010) with permission.

From Figure 1-3, it is evident that heart rate can vary depending on age, but other factors such as physical fitness level, stress, or exercise can also affect heart rate. In addition, each cycle of the heart contains more health information than just the heart rate. The rhythm of the heart and the coordination of each of the cycles reveal whether

the heart is diseased or infected (McCann et al. 2005c). One method for identifying indicators of heart disease is to simply listen to the sounds the heart makes.

1.3.3 Classes of Heart Sounds

When listening to heart sounds, a clinician typically notes the following characteristics for abnormal heart sounds: location on the chest, timing, intensity, and the quality or shape of the sound (Blood 2007). While the patient lies down, a clinician will start auscultation over the aortic region, followed by the pulmonic, tricuspid, and mitral regions, as illustrated in Figure 1-4.

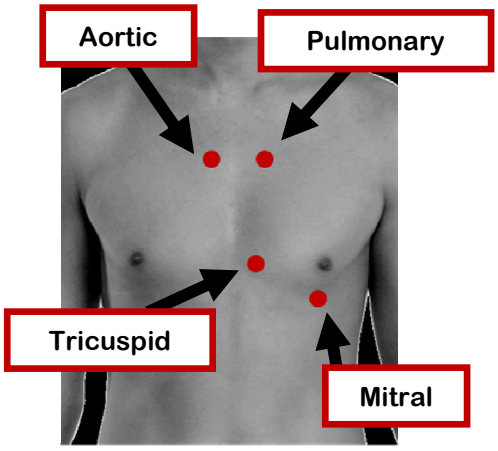


Figure 1-4. Locations on chest for auscultation specifically for listening to heart sounds. Adapted from Haggstrom (2009).

Cardiac sounds occur within a low frequency range from 2-150 Hz (Piinla 1995). The normal and abnormal heart sounds that can be heard by physicians during auscultation are described below. It is important to keep in mind that the audible presence of a certain class of heart sound can be normal for some ages but abnormal for others. Auscultation of the heart can reveal up to four heart sounds (Day et al 2009):

S1: (First heart sound) “Lub” sound of tricuspid and mitral valves closing and is heard the loudest at the apex of the heart (Douglas et al. 2009). Marks the beginning of systole when the ventricles start to contract, while the pulmonic and aortic valves open so that blood can flow to the lungs and the body.

S2: (Second heart sound) “Dub” sound of pulmonic and aortic valve closing and is loudest when heard at the base of the heart (Douglas et al. 2009). Marks the beginning of diastole when the heart relaxes (Joshi 1997). The atria contract, and the tricuspid and mitral valves open so that blood can start flowing to the ventricles (Meditec 2009). Sometimes there can be a “split S2” if there is a delay between the two valves closing. The length of the delay (or split) varies according to whether it is inspiration or expiration phase.

S3: Follows the S2 and occurs from the vibrations of the ventricle walls when the ventricles fill quickly with blood (McCann et al. 2005c).

S4: Second stage of ventricle filling and immediately precedes S1. Occurs from the vibrations of the valves and ventricle walls (McCann et al. 2005c).

Furthermore, a variety of terms exist to describe the abnormal sounds:

Gallop: Presence of an audible S3 or S4 sound (McCann et al. 2005c).

Pericardial friction rub: Sounds like sandpaper. Occurs when muscle is rubbing against a membrane and indicates pericarditis (Douglas et al. 2009).

Murmur: Irregular sound that indicates turbulent blood flow across a heart valve (i.e. back flow, increased flow constricted valve) (Douglas et al. 2009). Can occur

at any point in the cardiac cycle. Severity is graded on a scale from I to VI with VI being the loudest and most severe.

The strength and timing of these events and features are indicative of different conditions. For example, an early third heart sound (called S3) can be heard normally in children and adults under 40 years old (Meditec 2009). However for older adults, the presence of an S3 can mean problems with the left ventricle and an early sign of heart failure or infection (Mathew and Aggarwal 2008). Heart sounds can indicate other abnormal conditions such as hypertension (loud S2), pulmonic stenosis (ejection click after S1 and systolic murmur), and atrial septal defect (split S2 and systolic murmur) (Mathew and Aggarwal 2008, Brown et al. 2008).

1.3.4 Existing Research in Automated Heart Sound Analysis and Classification

Compared to respiratory sound analysis, the development of cardiac analysis systems has advanced more quickly (Piinla 1995). This difference is due partially to the fact that cardiac sounds are more audible and easier to record than lung sounds. Heart sounds have a narrower frequency range, come from a more limited location on the chest, and can be captured without excessive respiratory sound interference. When detected, abnormal heart beats can signify diseases such as dysrhythmia, valvular disease, stenosis, or pericarditis (Wang 2007).

In existing research, scientists have explored automated heart sound diagnosis using the phonocardiogram (PCG) signal, which uses a sensitive microphone to amplify

the vibrations of the heart and record them digitally or onto graph paper (Blood 2007). Liang et al. (1997) developed a segmentation algorithm that was based on separating the heart signal into four parts: S1, systole, S2, and diastole. The heart sounds were captured with an electronic stethoscope for a total of 37 recordings (515 heart cycles with normal heart sounds and murmurs) from children of age 0.4 to 14 years old. The algorithm first calculated the envelope of the signal from its normalized mean Shannon energy. The peaks in the resulting envelope signal that passed a threshold were considered S1 or S2 sounds. The systolic and diastolic intervals were estimated and extra peaks were discarded as signal artifacts if they didn't fit within a certain tolerance of the estimated durations. If an insufficient number of peaks were detected, then the threshold was iteratively lowered until more peaks were identified. Based on Liang's assumption that the lower and upper limits of a heart sound duration was from 20 to 120 ms, S1 and S2 sounds were distinguished from one another by finding another set of boundaries to describe the time duration of S1 and S2 individually. The algorithm, with 94.1% sensitivity and 98.8% positive predictivity, demonstrated the ability to successfully extract useful information from heart sound waveforms.

A duration-dependent hidden Markov model (DHMM) was shown to be another method to segment heart sounds (Schmidt et al. 2010). The inputs to the model were the duration parameter (estimated using autocorrelation analysis), extracted features, and envelope of the heart sound signal. The hidden Markov model, which had four states: S1, silent part of systole, S2, and silent part of diastole, was then used to automatically detect the locations of the S1 and S2 sounds in a heart sound waveform. Heart sounds were recorded for eight seconds with a Littmann electronic stethoscope

from 73 patients who were referred for a coronary arterioangiography. The DHMM was able identify to 890 S1 and S2 sounds out of 901 total (when manually segmented), for 98.8% sensitivity and 98.6% positive predictivity. Despite these good results, data from 15 patients (16% of the original 88 patients) were excluded because manual segmentation was not possible. Furthermore, there were errors in automated heart sound segmentation for 25% of the cases where the patients had arrhythmia due to variability in duration of the diastole and systole.

Malarvili et al. (2007) also analyzed heart sounds using the PCG, but incorporated information from the instantaneous energy of the ECG to enhance automated detection of the S1 and S2 sounds. There is a relationship between the timing of the signal characteristics of the ECG and heart sound signals, such that the end of the first peak in the ECG signal correlated with the peak in the heart sound for S1, while the end of the next peak in the ECG signal matched with the second heart sound. Malarvili et al. calculated the instantaneous energy of the ECG and PCG signals from the squared amplitude of the sum of the original signal *x(n)* and the Hilbert transform[29] of *x(n)*. Then they aligned the simultaneous ECG and PCG signals and were able to identify S1 and S2 sounds based on peaks in the instantaneous energy in the ECG. They tested the algorithm on a total 210 cardiac cycles from five normal and 10 abnormal patients, who were diagnosed with mitral regurgitation, mitral stenosis, and ventricular septal defect. They reported the method as an "effective algorithm," but quantitative results were not provided. Furthermore, the ECG signal was necessary in order to identify the S1 and S2 sounds.

[29] Used for exploring system "stability, linearity, and causality" and is calculated as the convolution of x[n] and h[n] where h[n] is 0 for n even and 2/(πn) for n odd (Worden and Tomlinson 2001, Tsui, J., 2004).

Clinicians are familiar with the observation that patients with morphologically similar ECG and heart sound waveforms likely indicate the same cardiovascular disease diagnosis (Wang et al. 2007). As a result, Wang et al. developed methods to extract shape and morphological information from ECG and heart sound signals in order to find disease similarity among patients. Then they used dynamic time warping to align a patient's ECG and acoustic data to signals from other patients with known diseases. They checked the alignment of corners in the signal (that define the morphology of a QRS complex for example), assigned a matching score based on mean square error, and then ranked the scores to identify the best match. Wang et al. tested the shape matching algorithm on 170 sample recordings from reference CDs, Boston Children's Hospital, University of Washington, UCLA, and UCSD Medical Schools. While the article illustrated examples of the algorithm applied to several signals, the performance of the algorithm on the entire collection of recordings was not reported in the article. The ambiguity of the results section suggests that this research is still evolving and not yet ready for use as a decision support tool in a clinical setting.

Syed (2003) performed similar work by using a 2-lead ECG and an electronic stethoscope to detect murmurs due to mitral valve prolapse (MVP). The motivation behind this automated auscultation system was the high false positive rate (80%) for patients referred to cardiologists for MVP using chest auscultation. Detection of MVP from heart sounds depends on higher frequencies being present in the second half of the systole. Syed (2003) developed an algorithm that first segmented the heart beats, selected out the beats that were noise-free, and then focused on murmurs in the systole (within the interval between the QRS complex and S2). Syed split the signal into 16

different frequency bands, followed by frequency band aggregation to find relevant frequency components of each heart beat. He also used beat averaging to generate a "typical" heart beat representative of the patient, and then performed frequency band-specific thresholding to decide whether the patient had MVP. From testing 51 patients where 21 had MVP, the number of false positives was reduced from 80% to 10% (where the 80% level of false positives came from diagnoses by primary care practitioners). Despite the promising results from Syed's experiments, the system is not as applicable in the context of resource-poor regions because Syed's system was intended for use in a primary care physician's office in a developed country and requires a good quality ECG and electronic stethoscope.

To address the challenge of scarce medical equipment and the associated consumables (such as the ECG and electrodes) in developing countries, Jin et al. (2008) are developing a low-cost ECG device to be plugged into mobile phones for early detection of heart problems including cardiovascular disease. Their "HeartToGo" project involves a wireless ECG sensor, and GPS device to collect and transmit location-specific ECG data to mobile phones for real-time processing. They plan to incorporate a person's medical history, physiological conditions, level of activity, and heart rate patterns to output a cardiovascular health assessment for an individual. However, the system is still being developed, and it is unclear if they will be able to overcome the issues associated with lead selection and consumables. ECG equipment is normally costly because of the numerous sensor leads and need for conductive electrode patches. With low-cost ECG sensors, there is an inevitable concern that the data being collected from cheap sensors could be distorted. Such systems will also

require reliable supply chain infrastructure to deliver the diagnostic peripherals, and expert knowledge to use the equipment. It is also difficult to remove noise sources such as baseline wander and muscle artifact, which can trigger a high number of false positives in any detection algorithm. The approach taken in this thesis avoids creating a low-cost version of a normally very expensive technology such as the ECG. Instead, by making use of existing equipment and supply chains (for mobile phones and hands free kits), it may be possible to perform clinical analysis of patient data.

In the commercial space, Bentley developed the iStethoscope Pro, which is an iPhone application that uses the iPhone 3G as a stethoscope to listen to heart sounds (Bentley 2009). The user presses the base of the phone (where the microphone is located) over the apex of the heart against bare skin. The phone application shows a phonocardiograph and spectrogram of the heart sound with many configurable features to modify sampling rate, audio delay, amplification level, and filter cut-off frequencies. On his website, Bentley (2009) claims that the tool is not intended for diagnostic purposes or substitution of real medical devices. There is however an option to email the recordings to Nordehn and Strunic, who created the iAuscultate (2006) application. The iAuscultate application for the iPhone is a training tool for heart auscultation based on repetitive listening of various types of heart sounds. Nordehn and Strunic are trying to create a large collection of heart sounds in order to develop automated heart sound classification systems. There is no standardized protocol on how to collect the data, and no publications exist yet to describe the quality of data being recorded. Users are allowed to place the microphone anywhere on the chest and can use any other type of microphone with the iStethoscope Pro application as long as it is noted during

submission (iAuscultate 2006). Perhaps more importantly, there are no indications that the system is able to deal with ambient noise or work across different hardware (which have different frequency responses). Finally, no discussions of compression or encoding exist for the application.

1.3.5 Fetal Heart Rate Monitoring

In addition to automated diagnostics for adult heart sounds, automated fetal heart rate (FHR) analysis from auscultation is also promising. Fetal heart sounds are very low frequency ranging from 0.1 Hz to 70 Hz (Mittra 2007). They are characterized primarily into S1 (20 - 40 Hz) and S2 (50 - 70 Hz) sounds. The signal is often weak because of the fetus' small heart valves and because of the many layers of tissue that the sounds must pass through to reach the surface of the mother's abdomen (Mittra 2007).

Earlier detection of fetal stress through audio heart rate analysis may provide appropriate interventions to reduce child and maternal mortality. Currently, fetal monitoring in developed countries heavily involves the use of ultrasound, but the technology is very costly and extremely problematic (because of the training required) to widely distribute in a developing country. In the U.S., there are many commercially available fetal Doppler ultrasound devices for personal use at home to monitor fetuses starting at 10-12 weeks of pregnancy (BabyBeat 2009). These devices range in level of sophistication but can include a display, recorder, audio speaker, and 3MHz probe. However, this type of ultrasound technology is extremely costly. A single device can cost $450 - $600, such as the product offered by babybeat.com (2009). The robustness of such devices is unknown at this time. Because of the steep cost of commercially

available ultrasound devices for professional clinical settings ($10,000 - $50,000), there is ongoing research to modify ultrasound technology or find alternatives to fetal monitoring with equipment that can fit the resource constraints of developing nations.

For example, Richard et al. (2008) and Tan (2009) are attempting to create low-cost and portable Doppler ultrasound devices[30] that can be plugged into a Windows Mobile phone for an application such as screening pregnant women in rural communities. Tan's project is ongoing, and the end goal is for mothers to be able to track and record the FHR of their infant over long periods of time while possibly transmitting the data to a specialist.

Another example of a FHR monitor, again Doppler-based, is the one designed by Hutchinson (2009) to work off the electricity grid. The ultrasound device is powered by a hand crank. A healthcare worker can crank the device for 1 minute to gain 10 minutes of operation time and then detect fetal distress in the case of an abnormal heart rate.

A recent advance in fetal monitoring comes from the Manoharbhai Patel Institute of Engineering and Technology in India (Mittra 2009). Mittra developed a two-microphone system to monitor abnormal disturbances in FHR, to be used with a pregnant woman when she is at a quiescent time just before falling asleep. During sleep, the mother's blood pressure and heart rate decrease naturally, but may cause an abnormality to occur in the FHR. One microphone is attached to the mother's abdomen for the fetal heart sound, while the other microphone is placed away from the body to capture ambient noise. The two audio signals are fed as input to a PC (located at the bedside) to analyze the data. The second (ambient signal) is used as an adaptive filter to remove exogenous noise sources from the fetal microphone. Unfortunately, this does

[30] Commercially available through Interson Inc.

not address the issues of endogenous noise sources such as placental blood flow, gastric activity, and the mother's breathing sounds. If abnormalities in the FHR are detected, the computer emits a warning signal, sends the audio data to a doctor, and notifies medical emergency officials through the Internet. An earlier prototype version of this system (without the bedside monitoring laptop application) was tested on 20 pregnant women between the 36^{th} and 40^{th} week of pregnancy. The results demonstrated 97.96% accuracy in FHR estimation from fetal heart sounds recorded with their monitoring system compared to simultaneous recordings from an ultrasound device (Mittra et al. 2007). The system as a bedside monitoring application however is yet to be tested to determine whether there are any clinical implications for improved health.

The fetal monitoring approach based on heart sound acoustics appears to be promising, although significant equipment is still required to perform the screening and the system could be much cheaper than ultrasound if the laptop was replaced by some low-cost digital signal processing hardware. Previous studies have demonstrated that clinically relevant information can be extracted after processing fetal heart sound signals. For example, irregular fetal heart rate trends can signify oxygen deficiency in the fetal blood (Moghavvemi et al. 2003). Using a PCG and stethoscope on the pregnant woman's abdomen, Moghavvemi et al. observed fetal heart rate, heart sounds, breathing, and movement of the fetus. Endogenous noises such as the maternal heart beat and breathing can be addressed by recording vital signs simultaneously with another acoustic microphone or ECG (for example), and then employing an adaptive real-time filter to attenuate such signals from the fetal PCG

signal. The recording system developed by them included a modified electronic stethoscope to capture mother and fetal PCG signals. The data was then displayed in real time on a personal laptop as a time series waveform alongside a power spectrum of the data. The software application had signal processing capabilities such as digital filtering for noise reduction. The users could also detect abnormal heart beats through visual interpretation such as a split in S1 sounds or the identification of heart sounds that would have normally been covered up by loud noise. Despite the development of a non-invasive data recording and display tool, their system was tested only on healthy patients and the demographics of the group are unknown. Their work was limited in the lack of results reported on the performance of the system for unhealthy pregnant women, but would likely be an area for them to explore next. Furthermore, the combination of a custom electronic stethoscope device and laptop workstation would be difficult to scale in a developing nation.

As discussed in section 1.2.4 on lung sound classification, the key to widespread adoption of an automated heart/lung sound diagnostics tool is to build the tool on a cheap and readily available platform (such as the mobile phone) with either no peripheral devices, or by using devices for which there is already an existing supply chain distribution infrastructure. This thesis therefore aims to develop a prototype for a flexible mobile diagnostic platform for heart and lung sounds.

Chapter 2

Heart Sound Analysis Using a Standard Database

2.1 Background and Data Sources

Although there have been several attempts to create low-cost diagnostic mobile-phone-based systems for use in resource-poor areas, almost all of them lack formal and peer-reviewed scientific validation on their diagnostic accuracy (as detailed in Chapter 1). Without formal and open validation, diagnostic devices have the potential to be more harmful than helpful. High false negative rates may not alter the status quo significantly, except for a possible placebo effect or unless the device replaces more meaningful diagnostic tests. However, high false positive rates have the potential to stress a healthcare system with unnecessary demand, expenses, and time loss (leading to significant income loss for those in resource-poor countries). More importantly, unnecessary medical intervention can detrimentally affect patient health outcomes through stress, admission to the ICU, risk of infection from surgeries, prescription of antibiotics (which can lead to population resistance), and administration of toxic procedures (including radiodiagnostics, radiotherapy and chemotherapy). Even in developed countries, there are no standards bodies that provide guidelines on the evaluation and usage of ad-hoc diagnostic devices. As a result, anyone can build a

diagnostic device and release it to a global market instantaneously through web downloads for little or no cost.

In order to develop an automated heart and lung sound diagnostic tool, the algorithms that are applied to the patient audio recordings must be validated. Scientific validation allows for prediction of the efficacy of a given screening method and enables users to make informed decisions about the adoption of the technology. An example of a database used for algorithm evaluation is the expert-annotated MIT-BIH Arrhythmia Database - available for free from PhysioNet (Goldberger 2000). According to the ANSI and FDA guidelines (ANSI/AAMI EC38 and EC57), all arrhythmia analysis algorithms intended for the US market must be tested on the MIT-BIH database (AAMI 2003, 2007). In a similar manner for heart and lung sound analysis algorithms, an analogous expert-annotated “gold standard” database is required to allow the global community to determine and compare the performance of their algorithms.

This chapter focuses on a robust heart rate estimation algorithm for acoustic data, while future research will explore more sophisticated algorithms for automated detection of abnormal heart conditions (such as murmurs) from acoustic data. The accuracy of the algorithm for heart sounds was determined by using simultaneously recorded ECG signals as reference. The ECG records electrical activity in the heart by measuring the electric potential difference between electrodes placed on different locations on the surface of the body (Ashley 2004). A labeled ECG trace (with the P wave, QRS complex, and T wave) is shown in Figure 2-1. The QRS complex represents depolarization of the ventricles, which causes the heart muscle to contract for one heartbeat (Ashley 2004). The heart rate - defined as the number of heartbeats in a

minute - can be determined in the ECG trace from the RR intervals.[31] Similarly, in acoustic data, each S1 corresponds to a single heartbeat. Thus, the heart rate can be computed from the interval between one S1 heart sound to the next. When ECG and acoustic sounds are recorded and displayed simultaneously (see Figure 2-2), the QRS complex in the ECG is seen to align with the S1 heart sound for each heart beat. The performance of the heart rate estimation algorithm for heart sounds can then be quantified by comparing detection of S1 sounds in acoustic data against QRS complex detection in ECG signals.

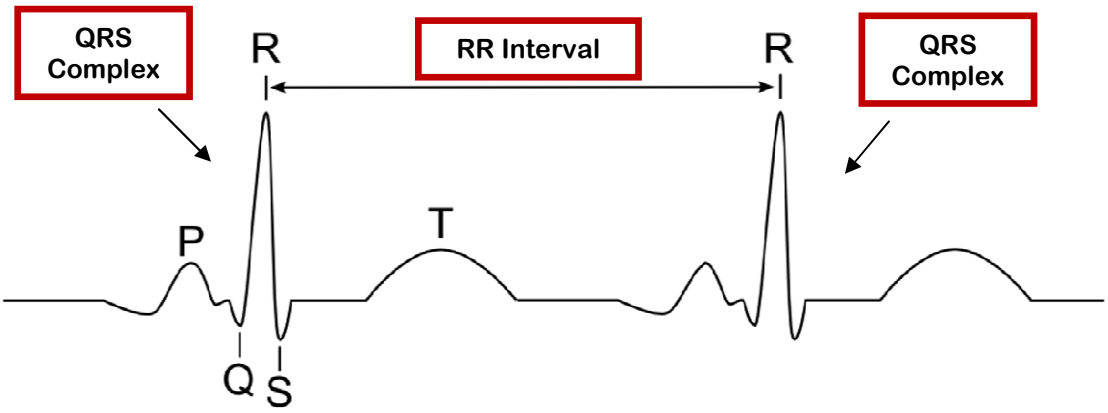


Figure 2-1. QRS Complex and RR interval on ECG waveform. Adapted from Burke (2008).

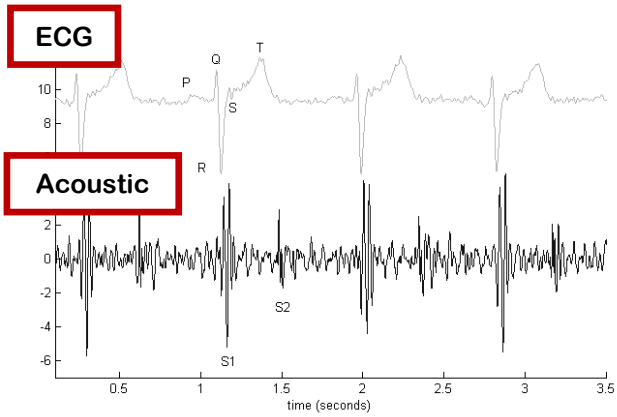


Figure 2-2. ECG and Associated Heart Sounds for test subject.

[31] An RR interval is the distance from one R peak to another and represents the number of seconds for a single heart beat to occur.

The heart sound database selected for use in this study was created at the Massachusetts General Hospital, as part of Syed's (2003) work to develop an automated software application for detection of mitral valve prolapse (MVP) in patients. The database consisted of cardiac heart sounds and ECG data from 51 patients (30 healthy individuals and 21 with MVP), who were referred for an echocardiogram. Syed's application was developed for use in a doctor's office, where a diagnosis was generated based on patient signals from an electronic stethoscope and two-lead ECG. Heart sounds were recorded with a Master Elite Plus Welch Allyn Meditron electronic stethoscope for approximately 30 seconds on each subject. The stethoscope had a frequency response of 20 Hz to 20 kHz. Recordings were performed with the stethoscope Bell setting, which applied a bandpass filter from 20 Hz to 420 Hz. The heart sounds were recorded from subjects in a supine position on the tricuspid region of the chest (see Figure 1-4). Files were stored as WAV files without compression at a 44.1 kHz sampling rate with 16-bit quantization. Out of 123 recordings, 27 clean ECG (with associated audio recordings of undetermined quality) from healthy adults were chosen for analysis. An example patient record containing an ECG and acoustic signal is displayed in Figure 2-2.

2.2 Heart Rate Estimation Methods

2.2.1 Pan and Tompkins' QRS Detector

The Pan and Tompkins' (1985) algorithm for real-time detection of QRS complexes in ECG traces was the basis of the heart rate extraction methods discussed

in this chapter. The Pan and Tompkins' (1985) algorithm detected QRS complexes (or heart beats) using a series of cascaded filters, followed by a segmentation and search procedure. The specific stages of the algorithm are shown in the Figure 2-3 block diagram.

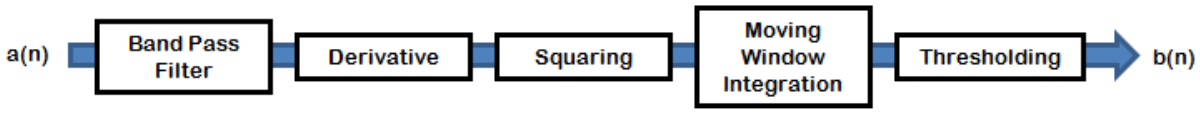


Figure 2-3. Block Diagram of Pan and Tompkins' (1985) QRS detector.
a(n) is the raw ECG signal and b(n) is the output of the QRS detector algorithm.

1. The first stage applied a **band pass filter** to the original raw ECG signal for noise reduction. The filter only allowed signal content between 5-11Hz to pass through, which was the ideal frequency range to maximize QRS energy (Pan and Tompkins 1985). Pan and Tompkins implemented the band pass filter using a low pass filter followed by a high pass one. The low pass filter had a cutoff frequency of 11Hz, delay of 6 samples, and gain of 36. Equation 2.1 is the low pass filter difference equation, where *a(n)* is the input signal, *lpf(n)* is the output of the filter, and *T* is the sampling period.

$$lpf(nT) = 2lpf(nT-T) - lpf(nT-2T) + a(nT) - 2a(nT-6T) + a(nT-12T) \qquad (2.1)$$

The subsequent high pass filter had a cutoff frequency of 5Hz, delay of 16 samples, and gain of 32. The difference equation of the filter (see Equation 2.2) has input *lpf(n),* and output *hpf(n)*.

$$hpf(nT) = 32lpf(nT-16T) - [hpf(nT-T) + lpf(nT) - lpf(nT-32T)] \qquad (2.2)$$

2. A five-point **derivative** of the signal was then taken, adding a delay of 2 samples. This step was used to obtain information about the slope of the QRS complex and was calculated according to Equation 2.3 to produce *deriv(nT)*.

$$deriv(nT) = (1/8T)[-hpf(nT-2T) - 2hpf(nT-T) + 2hpf(nT+T) + hpf(nT+2T)] \quad (2.3)$$

3. The resulting signal was **squared** (see Equation 2.4) to accentuate the rapidly changing high-frequency segments of the signal, where the QRS complexes resided.

$$sq(nT) = [deriv(nT)]^2 \quad (2.4)$$

4. Next, a **moving integration window** was applied across the signal to create a series of energy packets over each QRS complex. Note that the integration window should have a width that is roughly the size of the widest QRS complex. Too short a window decreased the signal-to-noise ratio, and too long of a window may have resulted in the merging of energy packets. For a signal sampled at 200Hz, Pan and Tompkins (1985) empirically determined the width of the window to be 30 samples or 150 milliseconds wide. The steps for calculation are described in Equation 2.5 with input *sq(n)*, output *integ(n),* and an integration window of *N* samples.

$$integ(nT) = (1/N)[sq(nT-(N-1)T) + sq(nT-(N-2)T) + ... + sq(nT)] \quad (2.5)$$

5. The QRS complexes yielded the highest signal amplitude points in the original ECG within the segments corresponding to the high energy regions. To determine whether a local maximum in the integrated waveform was a signal or noise peak, an **adaptive threshold** was used. To be classified as a signal peak (indicating a QRS complex), the peak had to be greater than threshold T_1. This threshold was continually updated with each new peak discovered. Another threshold T_2 was used in the case where a QRS complex could not be found in an expected period of time (RR interval). The algorithm proceeded to search backwards to find a peak that at least exceeded the lower threshold T_2. The thresholds were calculated with equations 2.6 to 2.9, where noise

peak *NP* or signal peak *SP* variables were updated if *P* was a noise or signal peak, respectively (Pan and Tompkins 1985).

$$T_1 = NP + .25(SP - NP) \quad (2.6)$$

$$T_2 = (1/2)(T_1) \quad (2.7)$$

$$NP = 0.125 * P + 0.875 * NP \quad (2.8)$$

$$SP = 0.125 * P + 0.875 * SP \quad (2.9)$$

In order to determine whether a QRS complex was expected within a certain period of time, a running **average of the RR intervals** was maintained. To compensate for rapidly changing heart rates, two averages were stored. The first value RR_{avg1} was the average of the eight most recently detected RR intervals. RR_{avg2} was similar, except that it only took into account the RR intervals that fell between 92% to 116% of the previous RR_{avg2}. If a QRS complex was not detected within an upper bound limit of the average RR interval (equal to 116% of RR_{avg2}) then the algorithm searched backwards for a peak. RR_{avg1} was maintained so that the algorithm could still detect peaks properly even with a highly variable heart rate.

Pan and Tompkins also built in a 200 ms refractory period to reduce over-sensitivity to prevent the QRS detector from re-triggering directly following any detected QRS complex. This algorithm was evaluated on the MIT-BIH Arrhythmia Database, and heartbeats were detected with a sensitivity of 99.68% and positive predictivity of 99.63%.

An open-source MATLAB implementation of a batch QRS detector, based on a modified version of the Pan and Tompkins' algorithm, was written by Clifford (2002). One of the most important differences introduced to the algorithm was the decision rule for determining a signal peak. The median peak height of the five most recently

identified QRS complexes was stored, so that the threshold (for a new peak to be considered a signal peak) could be computed as 20% of the median peak height. If a peak passed this threshold, the algorithm scanned a window of samples in the associated energy packet whose values were at least 90% of the max peak height in that energy packet. When applied to the MIT-BIH database, Clifford's QRS detection algorithm produced similar results to Pan and Tompkins in beat detection with 99.33% sensitivity and 99.06% positive predictivity. Clifford's peak detector for ECG data was adapted in this research for the purpose of extracting heart rate from ECG (see Section 2.2.2) and acoustic data (see Section 2.2.3).

2.2.2 Heart Rate Estimation from ECG

The detection of QRS complexes, based on Pan and Tompkins' method and Clifford's implementation described in Section 2.1.1, was used for heart rate estimation from ECG signals. Figure 2-4 shows the algorithm in a block diagram format.

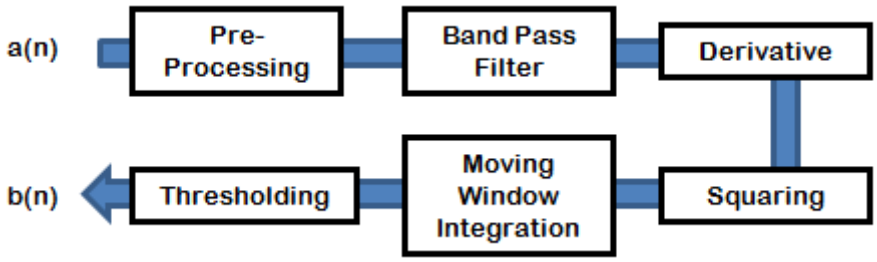


Figure 2-4. Block Diagram of modified Pan and Tompkins' (1985) QRS detector for heart rate estimation. a(n) is the raw ECG signal and b(n) is the output of the QRS detector algorithm.

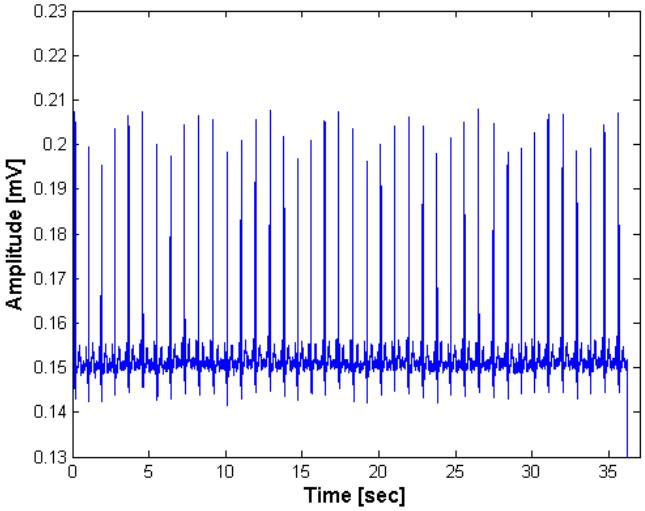


Figure 2-5. Original raw ECG signal from patient.

1. The **pre-processing** stage started with the raw ECG signal from Figure 2-5 and removed the first and last 500 samples of the signal, which were segments likely to contain noise.

The original ECG data from the database collected during Syed's research was sampled at 44.1 kHz, which had more precision than was needed for QRS detection. To improve processing speed, the ECG data was **downsampled** to 500 Hz. The resulting signal was verified against the original data to ensure that no salient features of the signal were lost.

The signal was then **normalized** by subtracting the mean and dividing by the standard deviation. The result of the pre-processing stage for a sample healthy patient is illustrated in Figure 2-6.

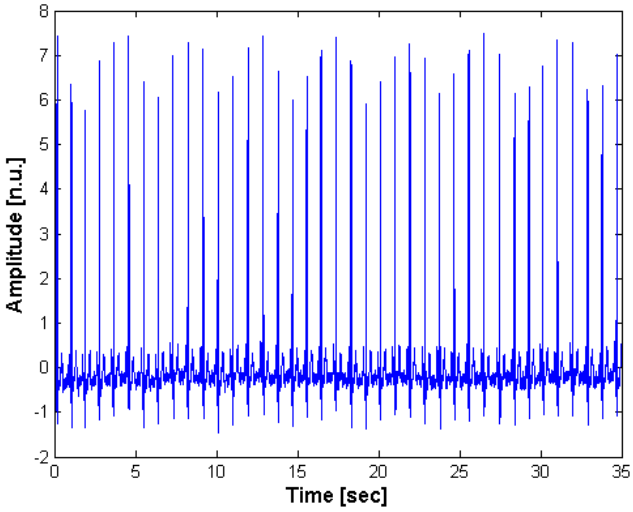


Figure 2-6. ECG signal after pre-processing stage.

2. Next, the signal passed through a 100-point FIR **bandpass** filter, whose magnitude and phase frequency responses are displayed in Figure 2-7. The low and high cutoff frequencies for the ECG filter were 2 Hz and 30 Hz, respectively, to reduce high frequency noise. The resulting vector was adjusted for sample delays from the filter (see Figure 2-8).

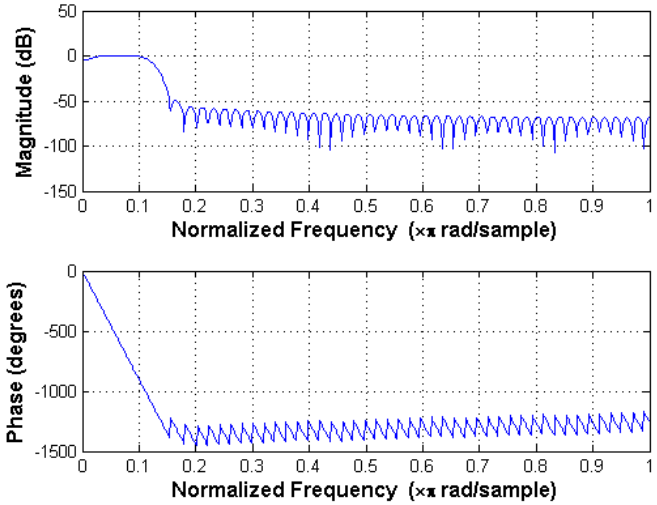


Figure 2-7. Magnitude and phase frequency response of band pass filter for ECG data.

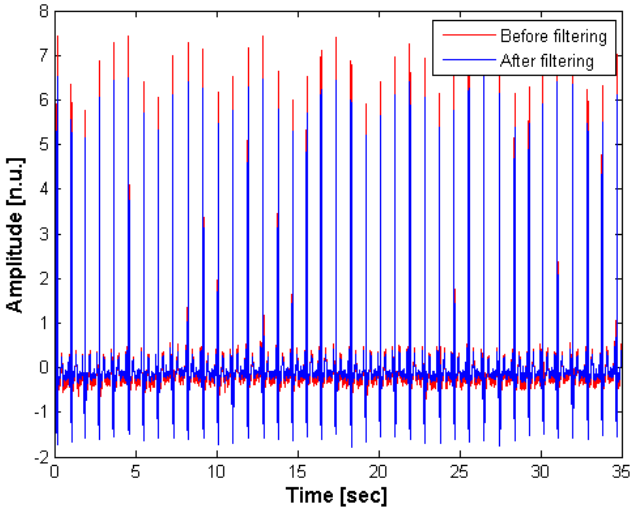


Figure 2-8. ECG signal before filtering (red) compared to after band pass filtering (blue).

3. The signal was then **differentiated, squared,** and **integrated** over a fixed-length window of 14 samples wide (or 28 milliseconds long) using a median filter. The output signal is shown in Figure 2-9.

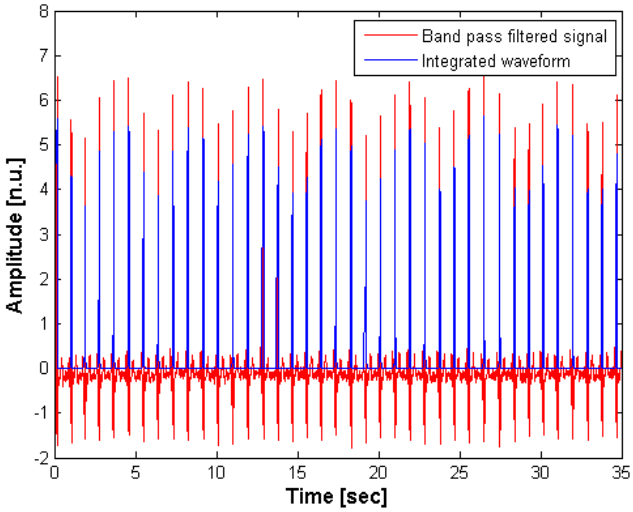


Figure 2-9. ECG signal after differentiation, squaring, and integration steps (blue) compared to signal before step 3 (red).

4. Within the integrated waveform, the search for peaks was confined to segments where distinct energy packets exceeded a peak triggering threshold of P_{thresh}. The P_{thresh} value was set at a fraction *t* of the largest sample amplitude found in s_{int} (see Equation

2.10), where s_{int} was the signal content from the 25th to 75th percentile of the ranked data from the integrated waveform. The algorithm searched each segment to find the R peak of the QRS complex, as Figure 2-10 demonstrates, and marked the R peak with a plus ('+') symbol.

$$P_{thresh} = t * \max(s_{\text{int}}) \tag{2.10}$$

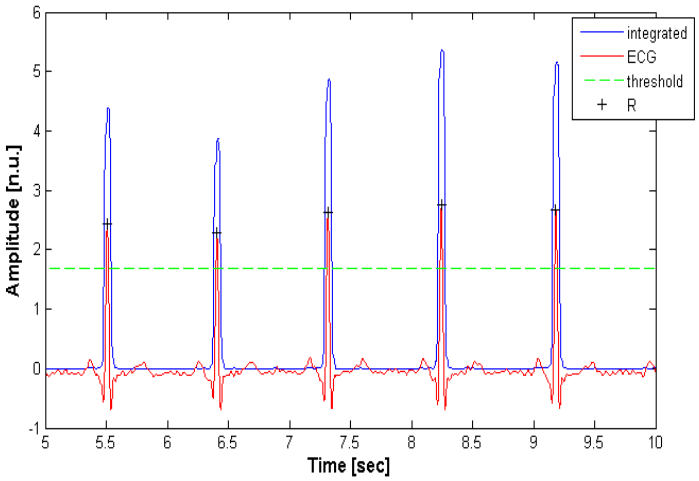


Figure 2-10. ECG R peak detection using integrated signal during five-second interval.

5. The algorithm produced two vectors: R_t containing the time indices of each R peak detected and $RR_{durations}$ with the length (in seconds) of each RR interval in the ECG trace. The instantaneous **heart rate** HR_i in beats per minute at the i^{th} beat was:

$$HR_i = \frac{60}{median(RR_{-4}, RR_{-3}, \ldots RR_3, RR_4)} \tag{2.11}$$

where RR_i was the RR interval between beat i and i+1. The median was used in Equation 2.11 to reduce variability and smooth out rapid changes in the RR intervals.

Since the ECG and heart sounds were recorded simultaneously for each patient, the location of the R peaks in the ECG were compared against the location of the S1 heart sounds on the same time axis. The identification of S1 sounds is described in Section 2.2.3.

2.2.3 Heart Rate Estimation from Acoustic Heart Sounds

S1 and S2 are the primary heart sounds and were introduced in Section 1.3.3 of this thesis. S1 marks the beginning of systole, where the tricuspid and mitral valves suddenly close so that blood can flow out of the ventricles and into the body and lungs (Taylor 2000). S2 marks the beginning of diastole, which is approximately twice as long as systole. The aortic and pulmonary valves close as the blood is pumped from the atria, through the tricuspid and mitral valves, and to the ventricles. The systole and diastole repeat in a cyclic manner for each heart beat (see Figure 2-11). The S3 and S4 sounds are sometimes heard in varying age groups, but may also indicate abnormal heart conditions. This research focuses on the detection of S1 and S2 heart sounds within acoustic heart signals in order to estimate heart rate as an initial proof-of-concept.

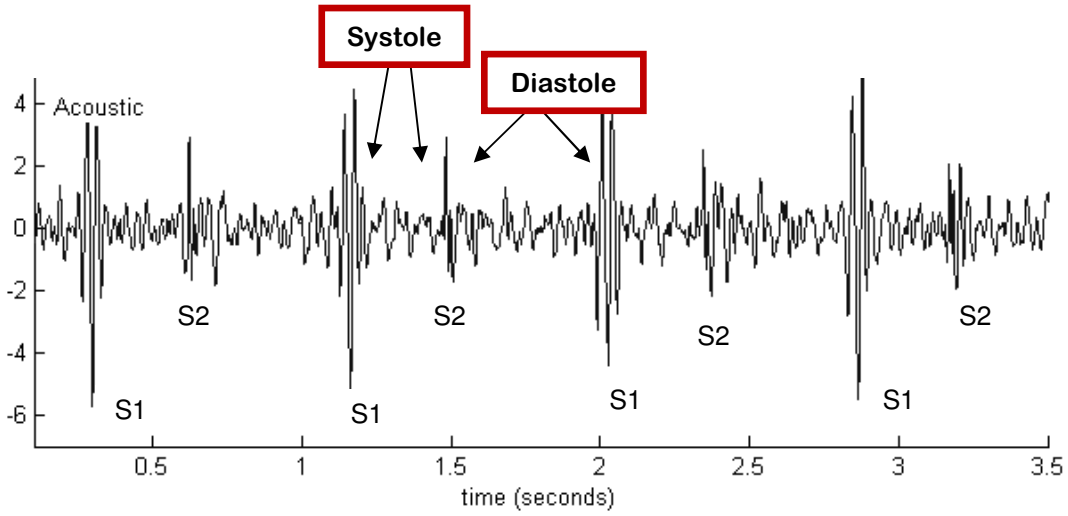


Figure 2-11. S1 and S2 heart sounds.

Due to commonalities between ECG and acoustic signals (high frequency spikes for each heart beat), a similar approach to heart rate estimation from ECG data is taken for heart rate estimation from acoustic data. The Pan and Tompkins' (1985) algorithm identified high frequency peaks as R waves belonging to QRS complexes in ECG

traces. Clifford's (2002) implementation of an R peak detector is adapted again, but instead for the detection of S1 and S2 heart sounds. The goal of the algorithm was to identify signal peaks in the acoustic waveform, distinguish S1 from S2 heart sounds, and calculate the heart rate based on intervals between S1 sounds. The block diagram of the algorithm, developed in close collaboration with Chen (2010) and Clifford, is presented in Figure 2-12.

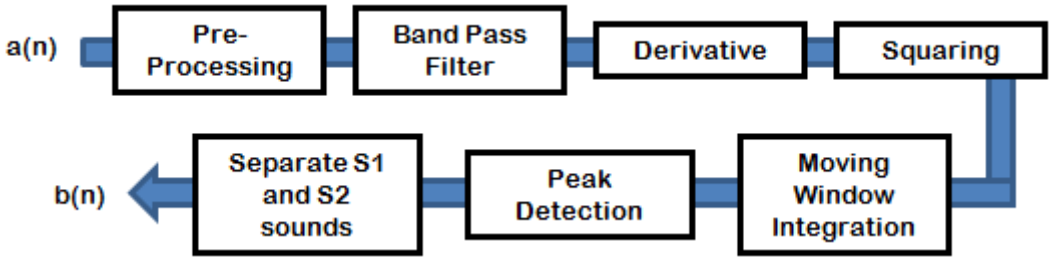


Figure 2-12. Block Diagram of modified Pan and Tompkins' (1985) QRS detector for heart rate estimation from acoustic data.

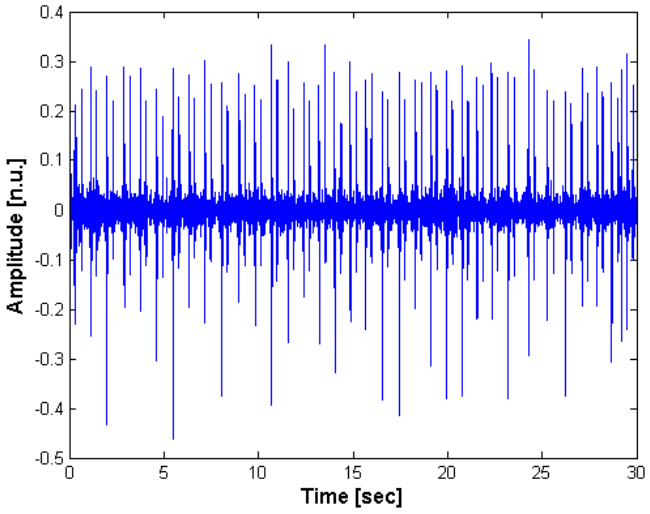


Figure 2-13. Original waveform of heart sounds recording with electronic stethoscope from healthy subject.

1. The **pre-processing** stage first **downsampled** the raw acoustic data (shown in Figure 2-13) from 44.1 kHz to 500 Hz and then **normalized** the signal. The new signal is shown in Figure 2-14.

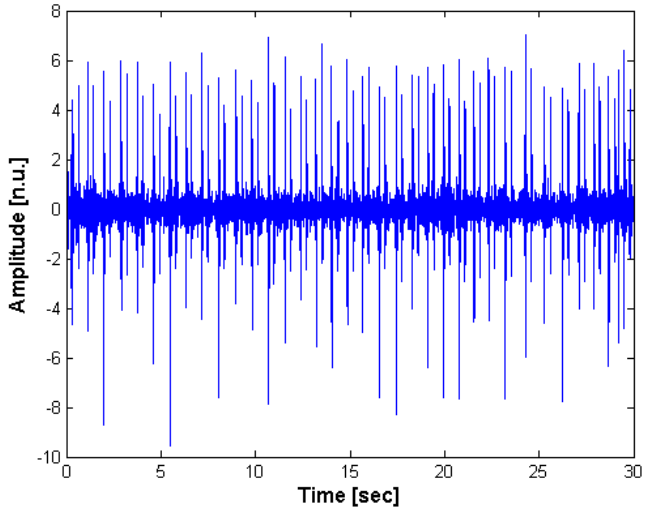


Figure 2-14. Audio recording after pre-processing stage.

2. The signal was then passed through a 100-point FIR **bandpass** filter with cutoff frequencies 5 Hz and 70 Hz. The magnitude and frequency response of the filter is graphed in Figure 2-15, while the output signal for a sample recording is shown in Figure 2-16.

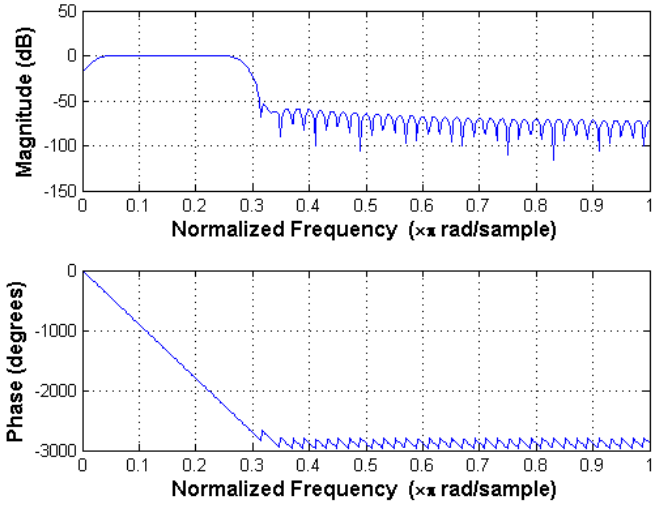


Figure 2-15. Magnitude and phase frequency response of band pass filter for audio data.

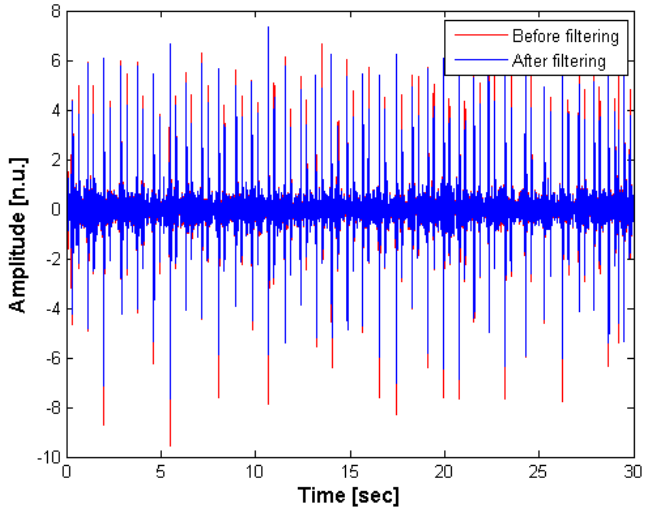


Figure 2-16. Audio signal before (red) and after band pass filtering (blue).

3. 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 29 samples wide (or 58 milliseconds long) using a median filter. Figure 2-17 shows the integrated waveform highlighted in blue over the previous ECG signal in red.

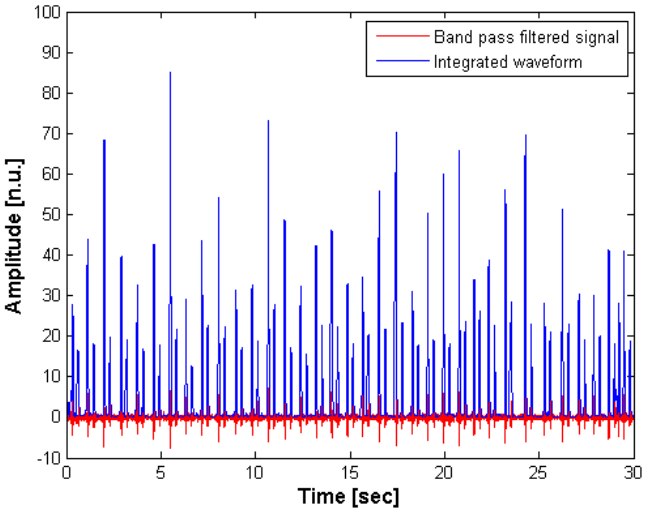


Figure 2-17. Audio signal before (red) and after differentiation, squaring, and integration (blue).

4. The resulting integrated quantity peaked in high energy areas, specifically during the S sounds (with potential to either be S1 or S2). These S sounds were labeled by finding the **local maxima** of the integrated signal (see Figure 2-18).

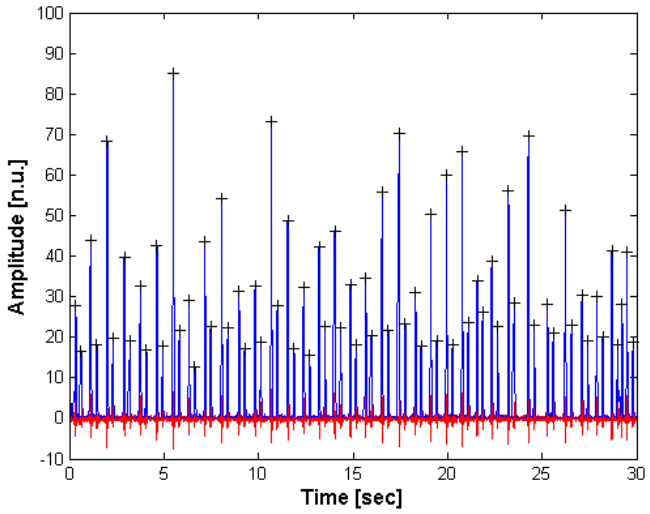


Figure 2-18. Audio signal after local maxima have been identified as S sounds (indicated with '+' symbol).

5. To **distinguish S1 from S2 sounds**, the time intervals (SS intervals) between detected local maxima were computed. A distribution of these SS intervals was created for each record, as evidenced in Figure 2-19, 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 (see Figure 2-6), the cluster of shorter SS intervals was designated as S1-S2, and the longer SS intervals as S2-S1. For each S1-S2 interval, an S1 sound at the start of that interval was labeled on the graph with a black plus symbol in Figure 2-20. For each S2-S1 interval, an S2 sound at the start of the interval was labeled on the graph with a green plus symbol (see Figure 2-20).

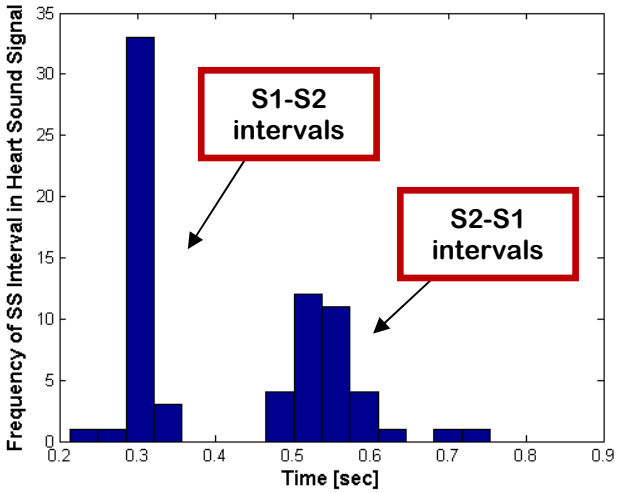


Figure 2-19. Empirical distribution of SS intervals for entire record of patient depicted in Figure 2-3. Adapted from Chen et al. (2010).

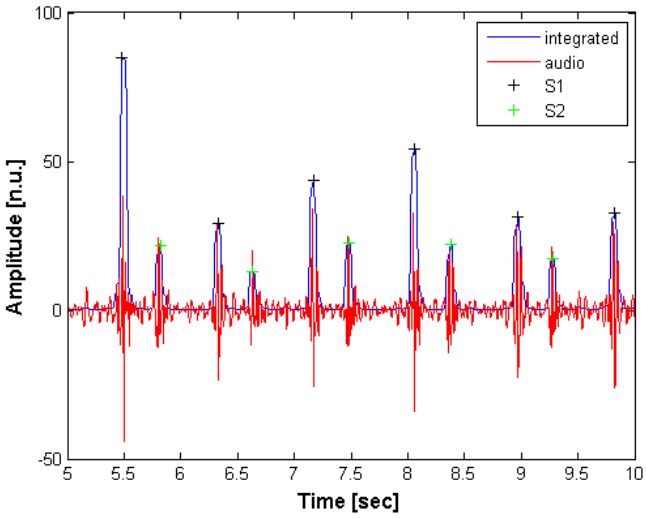


Figure 2-20. S1 and S2 heart sound detection using integrated signal derived from electronic stethoscope.

5. The algorithm produced vectors containing the time indices and amplitude of each S1 and S2 sound detected. The number of seconds for one heart beat to occur is measured from one S1 sound to the next, which is defined as an S1-S1 interval. Instantaneous **heart rate**, which is the number of heart beats in a minute, is then estimated from the median of S1-S1 intervals over nine beats. The heart rate HR_i at the

i^{th} beat is calculated in Equation 2.12, where SS_i is the S1-S1 interval between beat i and i+1.

$$HR_i = \frac{60}{median(SS_{-4}, SS_{-3}, \ldots SS_3, SS_4)} \tag{2.12}$$

2.3 Results

The accuracy of heart rate estimation based on S1 detection from audio data was compared to QRS (R peak) detection from simultaneously recorded ECG signals. Table 2-1 shows the difference between the actual heart rate (extracted from the ECG) compared to the estimated heart rate extracted from the acoustic signal, and relevant percentage error for each recording. The table also displays the number of correct, missed, and false heart beats detected by the algorithm for heart rate estimation from heart sounds. From the dataset in the table, the root mean square error (RMSE) is 23.2 bpm and the root mean square percentage error (RMPSE) is 3.13%.

In Table 2-1, there are two data outliers (circled in red) corresponding to recording file numbers 6 and 13. In recording 6, there were 29 beats that were falsely detected by the algorithm (meaning that they were not actual beats when the ECG trace was analyzed) - a 115% error in estimating the heart rate using the acoustic signal. Similarly for recording 13, there were 17 false beats detected which caused a 115% heart rate estimation error. A closer look at these two records revealed that the histogram of SS intervals could not be clearly divided into two Gaussian-like clusters for S1-S2 and S2-S1 intervals respectively. This ambiguity in distinguishing between the

intervals led to erroneous demarcation of too many S1 (or S2) sounds, as well as detection of too many beats.

ID	Actual HR (from ECG) in bpm	Estimated HR (from audio) in bpm	% Error	# Correct Beats	# Missed Beats	# False Beats
1	72.29	70.75	-0.02	36	0	2
2	61.48	62.11	0.01	30	1	4
3	74.81	74.63	0	23	3	2
4	72.82	72.46	0	42	1	0
5	90.91	91.32	0	54	1	9
6	79.37	170.45	1.15	42	1	29
7	65.65	66.82	0.02	40	0	7
8	72.64	72.46	0	36	6	3
9	69.44	69.61	0	40	1	0
10	54.25	54.55	0.01	28	5	4
11	68.65	68.97	0	40	0	1
12	89.02	89.55	0.01	51	0	8
13	68.81	147.78	1.15	33	3	17
14	67.57	68.51	0.01	39	1	2
15	71.09	71.09	0	35	2	5
16	70.01	69.45	-0.01	30	5	3
17	67.42	67.42	0	29	5	1
18	72.82	73.35	0.01	34	0	6
19	56.39	55.35	-0.02	26	2	0
20	64.24	64.24	0	32	1	2
21	71.86	70.42	-0.02	24	13	0
22	56.18	56.6	0.01	24	3	1
23	82.42	82.3	0	41	1	0
24	77.32	73.09	-0.05	18	21	1
25	63.97	66.53	0.04	30	2	5
26	50.42	50.25	0	26	0	2
27	65.79	68.65	0.04	32	1	6

Table 2-1. Results of heart rate estimation algorithm on heart sounds database.

When the two data outliers were removed, the percentage error in heart rate estimation for each record visually dropped to within 5% of the actual heart rate (in Figure 2-21). As a dataset, the RMSE of the recordings dropped to 1.3 bpm with a RMSPE of 0.19%. The best-fit linear trend with equation y = -0.00035x+.025 indicated

that there was no conclusive relationship between the patient's actual heart rate and the estimation error. Other factors were potentially more influential on estimation error such as the signal quality or the patient's health condition.

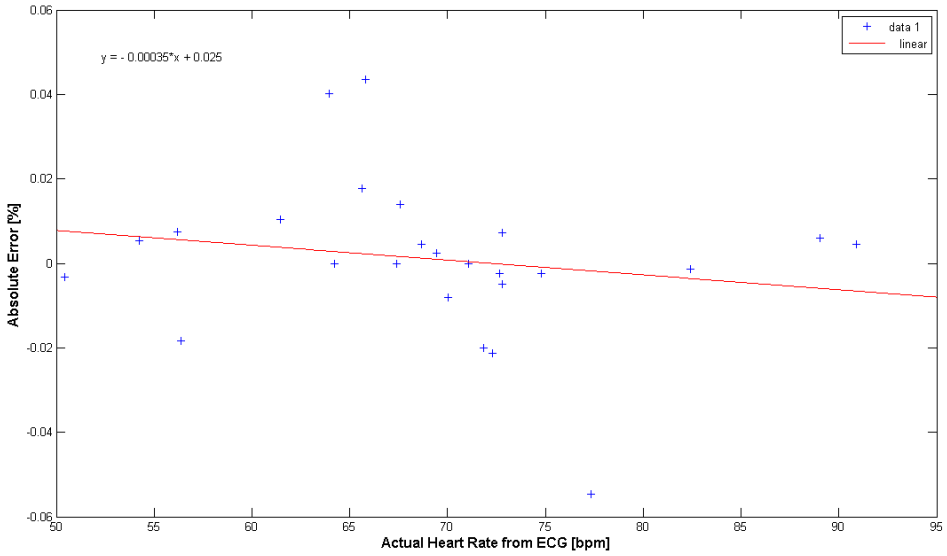


Figure 2-21. Data outliers removed. Percentage error of the estimation algorithm based on a patient's heart rate with fitted linear trend.

All records examined from the database had 100% R peak detection, totaling 994 beats. S1 detection was considered accurate if it was within 0.1 second of the corresponding R peak. The heart rate estimation algorithm detected 915 heart beats correctly, missed 79 beats, and falsely detected 120 beats.

2.4 Discussion

The heart rate estimator (and heart sound detector) proposed in this chapter is able to detect S1 sounds with a sensitivity of 92.1% and a positive predictivity of 88.4% when applied to a gold standard database of ECG and heart sounds. These results are

lower than the 99.9% levels reported for ECG beat detectors (Hamilton and Tompkins 1986) because the signal-to-noise ratio is lower in heart sound recordings than ECGs. However, the heart sound detection algorithm is sufficient to detect the majority of heart beats, and this is enough to evaluate the heart rate with reasonable accuracy. The key for this algorithm is to be able to detect whether the heart rate of the individual is within the acceptable range for healthy people. If the algorithm is providing abnormal heart rate values well beyond the expected range, then the clinician should conduct further investigation to determine the patient's risk for an abnormal heart condition. Detecting individual heart sounds is also the first step towards feature extraction for more complex disease detection. Further signal processing can now be performed to find characteristics such as shape morphology and exact timing of each heart sound (within the heart beat cycle). However, a labeled database will be required to develop such algorithms, and hence this is a future work consideration.

One interesting observation is the effect of respiration on the heart sound recordings. From Figures 2-18 and 2-20, the S1 and S2 complexes exhibit slow changes in average energy over a period of several seconds. This observation may be attributed to the fact that as a person breathes, the location of the heart can change, possibly moving towards and away from the microphone (depending on microphone location). Alternatively, the change in thoracic pressure from the movement of the diaphragm during breathing may also cause the changes in average energy over time in the heart sound waveform. Therefore, cardiac data may also be used to detect respiration rates in a similar manner to that in ECG analysis (Moody 1985).

An important issue involved in auscultation is the interference of noise - both from the ambient environment and from the patient's movements (i.e. during breathing). Being able to detect and remove noise artifacts will be an essential part to an automated analysis system such as the one proposed by Bhatikar et al. (2005). 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. As mentioned in Section 1.3.3, one method of noise reduction is to use a second microphone to record ambient noise to adjust an adaptive filter. Results in this domain have been promising but still require validation in terms of demonstrated improvements in clinical health outcomes (Mittra et al. 2009).

An important feature that will be required is the ability to inform the user when the recording location is providing sufficient signal quality for accurate and meaningful processing of the data. Li et al. (2008) and Nemati et al. (2009) described a signal quality index measure for ECG, blood pressure and respiration respectively, which could be applied to this framework for heart sound signals in future work.

Chapter 3

Lung Sound Analysis Using a Standard Database

3.1 Data Sources

As discussed in Chapter 2.1 for heart sounds, it is important to test any automated lung sound analysis algorithm on a standard database before use in a clinical setting, to provide users with an understanding of the accuracy of the algorithms, and domain of applicability. While the long-term goal of this research is to use signal processing to automatically detect abnormal lung sounds in patients as an indication of respiratory diseases, there is a scarcity of publicly available (and labeled) lung sound data. This means that there are an insufficient number and variety of recordings of adventitious lung sounds available to cover most of the types of problems one might encounter (as well as the heterogeneity of signals). A larger number and diversity of expert-labeled data are required to detect lung sound abnormalities. Hence, the work in this thesis aims to provide an initial proof-of-concept for basic lung sound analysis on audio recordings to perform automated respiratory rate estimation, a primary vital sign in respiratory-related illness diagnosis.

A collection of electronic stethoscope respiratory sound recordings, which were captured during the development of Yi's (2004) toolkit for lung sound analysis, was obtained from Dorothy Curtis of MIT CSAIL for the analysis in this chapter. This database contained 74 respiratory sound recordings from 13 adult subjects consisting of

6 females and 7 males. There were not consistent numbers of recordings for each subject, but 34% of the recording files were from female subjects. Further specifics on the demographics of the subjects were not provided (Yi 2004). The recordings were acquired with a Master Elite Plus Welch Allyn Meditron electronic stethoscope (similar to the one used for the database of heart sounds in Section 2.1). In addition, the data were saved in uncompressed WAV files (sampled at 22,050 Hz with a 16-bit amplitude resolution).

The 74 audio files had different durations ranging from 11 seconds to 100 seconds, with an average of 34 seconds. Since respiration rates or air flow graphs were not provided with the data, each file had to be manually labeled with the patient's actual respiratory rate in cpm. The labeling process was accomplished by listening to each WAV file and manually counting the respiratory cycles in the duration of the audio file. Respiratory rate (RR_{actual}) can be computed by Equation 3.1 with $N_{breaths}$ as the number of breath cycles counted and $T_{recording}$ as the length of the recording in seconds.

$$RR_{actual} = \frac{N_{breaths}}{T_{recording}} * 60 \tag{3.1}$$

There was approximately 42 minutes of data altogether, which includes a total of 1259 respiratory cycles. Table 3-1 provides the following information for each recording: i) file ID number (from 1 to 74), ii) subject ID number (from 1 to 13), iii) the manually-computed number of breath cycles in the recording, iv) duration of the recording in seconds, v) computed respiratory rate, and vi) gender of the subject.

File ID	Subject ID	Actual RR (cpm)	Actual Cycles	Actual Time (sec)	Recording Location	Gender
1	1	15.00	15.00	60.00	Trachea	M
2	1	17.00	9.00	32.00	Trachea	M
3	1	15.00	14.00	56.00	Trachea	M
4	1	15.00	15.00	60.00	Trachea	M
5	1	15.00	10.00	39.00	Trachea	M
6	2	14.00	13.00	56.00	Trachea	M
7	2	15.00	14.00	56.00	Trachea	M
8	2	11.00	9.00	49.00	Trachea	M
9	2	12.00	11.00	55.00	Trachea	M
10	2	14.00	12.00	52.00	Trachea	M
11	3	23.00	5.00	13.00	Trachea	M
12	3	23.00	5.00	13.00	Trachea	M
13	3	25.00	9.50	23.00	Trachea	M
14	3	25.00	12.00	29.00	Trachea	M
15	3	25.00	10.00	24.00	Trachea	M
16	3	25.00	12.00	29.00	Trachea	M
17	3	26.00	25.00	58.00	Trachea	M
18	3	27.00	23.00	52.00	Trachea	M
19	3	26.00	20.00	47.00	Trachea	M
20	3	29.00	20.50	42.00	Trachea	M
21	3	25.00	21.00	51.00	Trachea	M
22	4	16.00	6.00	22.00	Anterior chest	F
23	4	17.00	5.00	18.00	Anterior chest	F
24	5	18.00	8.00	27.00	Anterior chest	F
25	6	12.00	5.00	26.00	Anterior chest	F
26	6	17.00	5.50	19.00	Anterior chest	F
27	6	17.00	5.00	18.00	Anterior chest	F
28	6	14.00	5.00	22.00	Posterior chest	F
29	6	18.00	5.00	17.00	Posterior chest	F
30	6	14.00	5.00	21.00	Posterior chest	F
31	7	15.00	25.00	100.00	Trachea	M
32	7	14.00	8.00	35.00	Trachea	M
33	7	12.00	10.00	49.00	Trachea	M
34	7	13.00	12.00	55.00	Trachea	M
35	7	13.00	13.00	59.00	Trachea	M
36	8	15.00	6.00	24.00	Anterior chest	F
37	8	12.00	5.00	25.00	Anterior chest	F
38	8	15.00	5.50	22.00	Anterior chest	F
39	8	16.00	6.00	22.00	Anterior chest	F
40	8	13.00	5.00	24.00	Anterior chest	F
41	9	12.00	5.00	26.00	Anterior chest	M
42	9	13.00	5.00	23.00	Anterior chest	M
43	9	16.00	5.00	19.00	Anterior chest	M
44	9	15.00	5.00	20.00	Posterior chest	M
45	9	14.00	5.00	21.00	Posterior chest	M
46	9	16.00	5.00	19.00	Posterior chest	M
47	9	14.00	5.00	21.00	Posterior chest	M
48	9	14.00	5.00	21.00	Posterior chest	M
49	10	15.00	6.00	24.00	Trachea	M
50	10	15.00	6.00	24.00	Trachea	M
51	10	19.00	5.00	16.00	Trachea	M
52	10	18.00	6.00	20.00	Trachea	M
53	10	15.00	4.00	16.00	Trachea	M
54	10	17.00	9.00	31.00	Trachea	M
55	11	21.00	6.00	17.00	Anterior chest	F
56	11	22.00	5.50	15.00	Anterior chest	F
57	11	25.00	5.00	12.00	Anterior chest	F

58	11	19.00	6.00	19.00	Posterior chest	F
59	11	23.00	5.00	13.00	Posterior chest	F
60	11	21.00	5.00	14.00	Posterior chest	F
61	11	30.00	5.50	11.00	Posterior chest	F
62	11	21.00	5.00	14.00	Posterior chest	F
63	12	17.00	16.00	57.00	Trachea	M
64	12	14.00	14.00	60.00	Trachea	M
65	12	6.00	5.50	52.00	Trachea	M
66	12	8.00	7.00	56.00	Trachea	M
67	12	8.00	7.00	55.00	Trachea	M
68	12	13.00	12.00	56.00	Trachea	M
69	12	15.00	14.00	56.00	Trachea	M
70	12	17.00	15.00	54.00	Trachea	M
71	12	15.00	14.00	56.00	Trachea	M
72	13	12.00	9.00	45.00	Anterior chest	F
73	13	18.00	10.50	35.00	Anterior chest	F
74	13	18.00	10.00	34.00	Anterior chest	F

Table 3-1. Manually labeled respiratory rates for database of lung sounds from Yi (2004).

All signal analysis was implemented using MATLAB version 7.8 with the Signal Processing, Filter Design, and Statistics Toolboxes (MathWorks, Framingham, MA, 2009). Figure 3-1 illustrates a 15-second WAV recording in MATLAB (recording 53 from subject 10 in Table 3-1). Note the four respiratory cycles, each of which is comprised of two large energy bursts. Each energy burst, lasting around 1.5 seconds in this example, represents one phase of the respiratory cycle (either inspiration or expiration). The four cycles in the 15-second waveform equate to a respiration rate of 16 cpm for the individual.

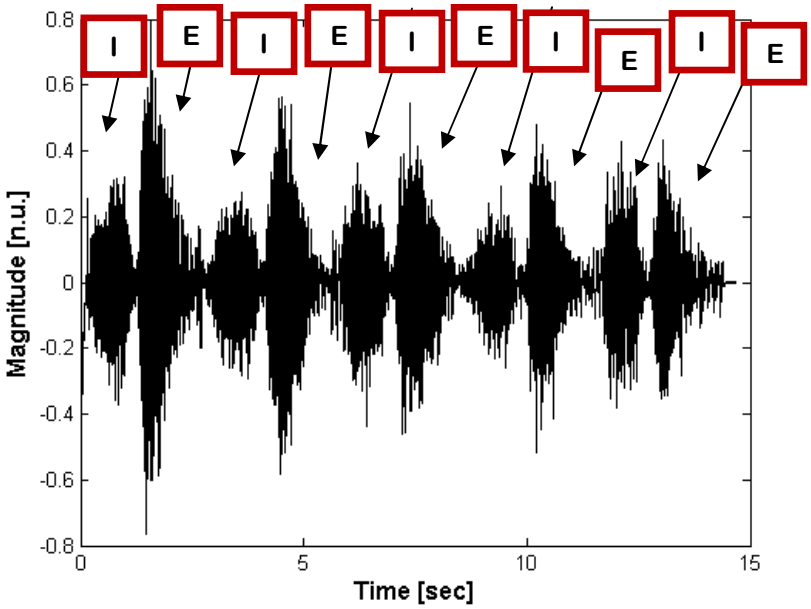


Figure 3-1. Time series plot of tracheal breath sound recording with labeled inspiration (I) and expiration (E) phases marked.

To automatically estimate respiration rate, there are essentially two approaches described in the literature: energy localization/peak detection of each phase of the respiration cycle or spectral analysis. Furthermore, these two methods are generally applied to the signal with varying levels of filtering, often in an attempt to extract the envelope and remove high frequency noise that could lead to false triggering. Three different algorithms, two of which were previously developed for respiration rate analysis, and one for heart rate analysis (described in Chapter 2) were evaluated on the lung sound database. These methods are now described together with adaptations required for the specifics of lung sound waveforms. Error profiles as a function of respiration rate were derived to test the hypothesis that: since each algorithm is expected to react differently to noise and that noise manifests in a different manner at different rates, varying estimation error rate profiles would be observed. A median voting system was also implemented to evaluate whether a 'product of experts' (in our case, the median of all the individual estimates) was a more robust method of

estimating respiratory rate than any individual algorithm. Finally, a previously developed signal quality index was evaluated to determine if thresholds could be set, below which recordings were untrustworthy and should not be used or require that the recordings be retaken.

3.2 Respiratory Rate Estimation Methods

3.2.1 Tracheal Breath Count Method

The first respiratory rate estimation method was based on the acoustic respiratory analysis method of Yi (2004). Her algorithm was tuned and validated on tracheal sounds from five healthy individuals - a portion of the data in the lung sound database used here. Her algorithm was tested on 60 respiratory phase examples from each subject for a total of 300 respiratory phases (compared to the 13 subjects and 1259 respiratory phases used in this thesis work). The data in Yi's work was only split into training and test sets for her respiration phase classification algorithm (to distinguish inspiration from expiration), but not for the respiratory rate estimator that is relevant to this section. Table 3-2 contains more information about the five subjects, whose average respiratory rate was 20 cpm with a standard deviation of 10 cpm, and the performance of the algorithm on each individual.

Subject Number	Gender	Respiratory Rate Range (cpm)	% Respiratory Phases Correctly Detected
1	F	6 - 16	100
2	M	15 - 16	93
3	M	12 - 17	95
4	M	24 - 29	98
5	M	11 - 15	100

Table 3-2. Subjects recorded for development of tracheal breath count method (Yi 2004).

The algorithm demonstrated an average accuracy of 97% ± 3% in terms of correctly detected respiratory phases. The original MATLAB implementation of this algorithm was made available for this research and is hereafter referred to as the Tracheal Breath Count (TBC) Method.

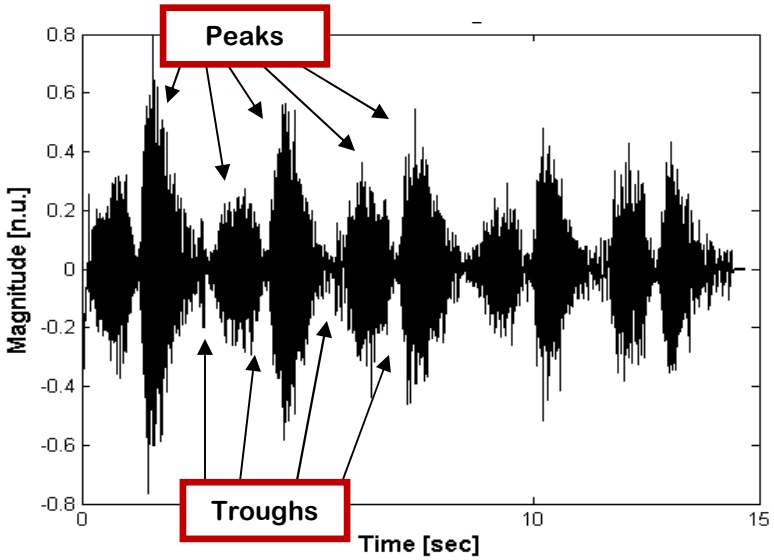


Figure 3-2. Peaks and troughs for each respiratory phase on tracheal breath sound.

The TBC method estimates respiration rate by determining the number of valid peaks in the envelope of the breath sound signal (Yi 2004). The method is based on the observation that each respiratory phase has a midpoint that is represented by a peak in the signal, and an endpoint that is represented by a trough (see Figures 3-1 and 3-2). The purpose of the algorithm was to extract the envelope of the acoustic signal such that its peaks and troughs were identifiable as single points. Then a respiratory phase was detected as a peak directly followed by a trough. Figure 3-3 shows a block diagram of the steps in the TBC method (Yi 2004) and can be summarized in 12 steps as follows:

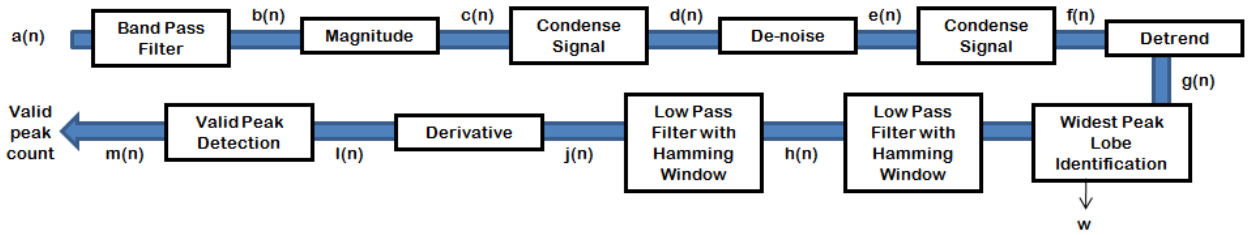


Figure 3-3. Block diagram of tracheal breath count algorithm (Yi 2004).

1. First an FIR **band-pass filter** with upper and lower cutoff frequencies of 100 Hz and 1300 Hz respectively was applied to the raw signal *a(n)* pictured in Figure 3-2. This frequency range corresponds to the band in which the majority of lung sound information has been observed to occur (Pasterkamp 2008). The filter was implemented by multiplying the truncated sinc function of the filter with a Blackman window to create a smooth frequency response curve in the frequency domain with a flat passband and good stopband attenuation (Smith 1997). The Blackman window function *bw(n)* was computed using Equation 3.2 for $|n| \leq Q$ (MathWorks 2009). In this case, the filter was of order $Q=f/12$ where f is the input signal's sampling frequency and the window length $N_B = Q+1$ (Yi 2004). An example of the resulting signal *b(n)* is seen in Figure 3-4.

$$bw(n) = 0.42 - 0.5\cos(2\Pi n / Q) + 0.08\cos(4\Pi n / Q) \tag{3.2}$$

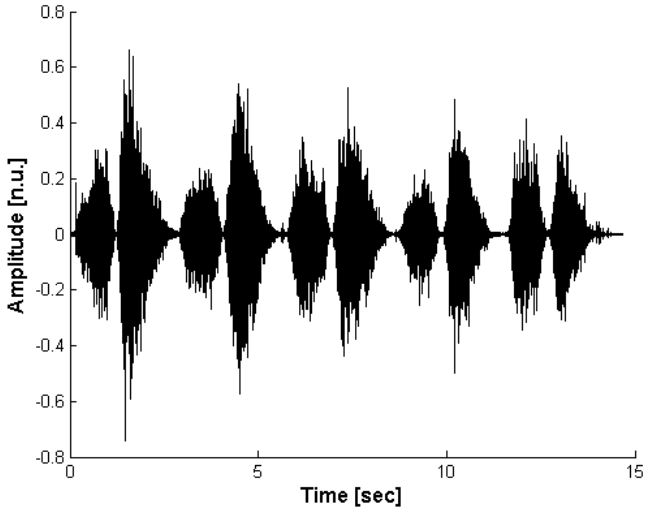


Figure 3-4. Band pass filtered lung sound signal b(n).

2. Then the **absolute magnitude**, *c(n)* = |*b(n)*|, of each point was taken, maintaining the peaks and troughs of the raw signal (see Figure 3-5).

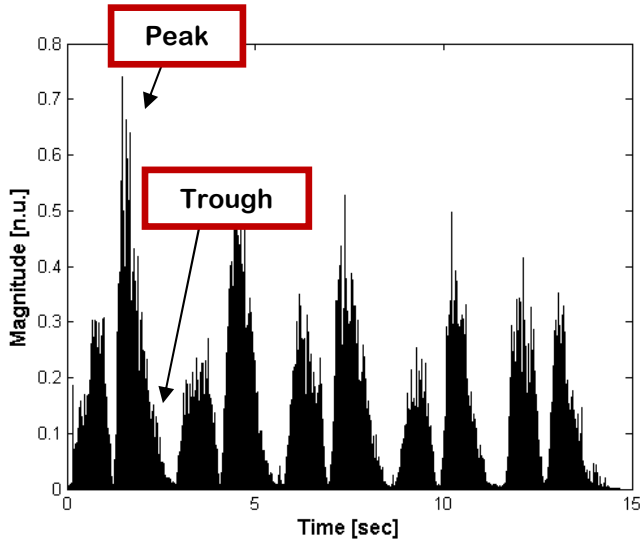


Figure 3-5. Absolute magnitude of band pass filtered lung sound signal c(n).

3. Only the envelope of the signal needed to be preserved for peak and trough detection, so computation time was reduced by generating a **downsampled signal** (Figure 3-6) from the lung sound signal. The downsampled signal d(n) was produced by calculating the mean of each non-overlapping contiguous 100-sample segment in *c(n).*

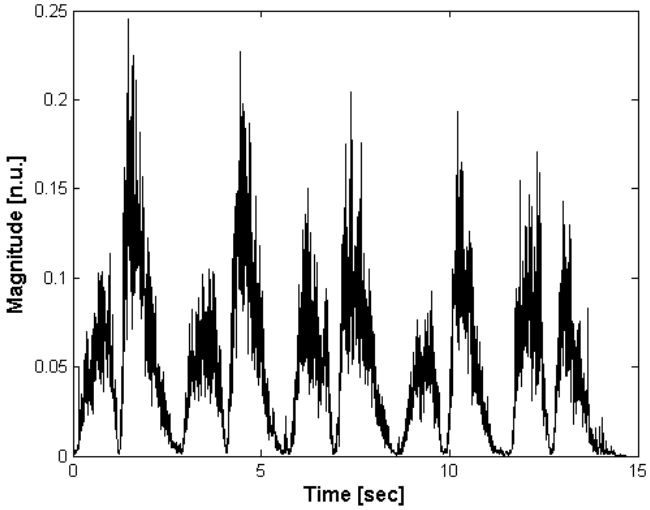


Figure 3-6. Downsampled lung sound signal d(n).

4. **Noise reduction** was achieved by a nonlinear 100-fold downsampling procedure from 8kHz to 80Hz to produce a new signal *e(n).* This involved segmenting *d(n*) into contiguous 80-sample segments. For a single segment within *d(n)*, a threshold was set to be equal to twice the mean of all samples within that segment. For each point *d(i)* at index *i* within that segment, if *d(i)* exceeded that threshold, then *e(i)* was set to the value of the previous point *e(i-1).* If *d(i)* was less than the threshold, then *e(i)* was set to the value of *d(i)* (See Equation 3-3). The new sampling frequency of *e(n)* as displayed in Figure 3-7 was 1/100th of the original signal *a(n)*'s sampling frequency.

$$e(i) = \begin{cases} 0 & if \quad d(i) > 2 * mean(segment_i) \,\&\, i = 0 \\ e(i-1) & if \quad d(i) > 2 * mean(segment_i) \\ d(i) & otherwise \end{cases} \tag{3.3}$$

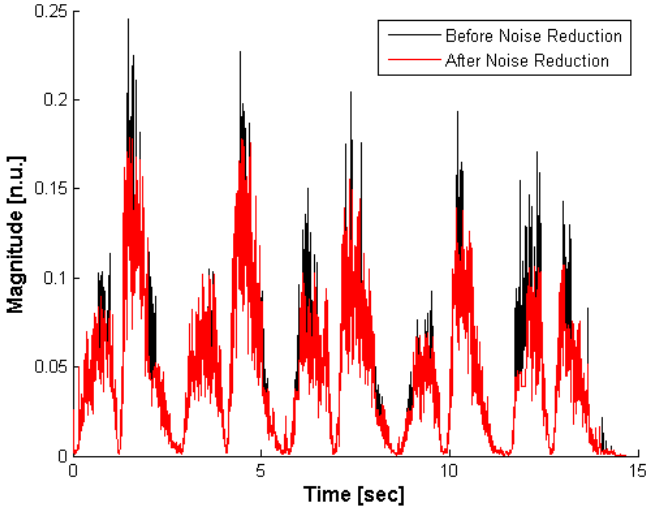


Figure 3-7. Before [d(n) in black] and after noise reduction [e(n) in red].

5. The signal *e(n)* was further **downsampled** to reduce computation load by defining *f(n)*, which was generated from extracting the mean of every contiguous non-overlapping 10-point segment. Figure 3-8 shows a plot of *f(n).*

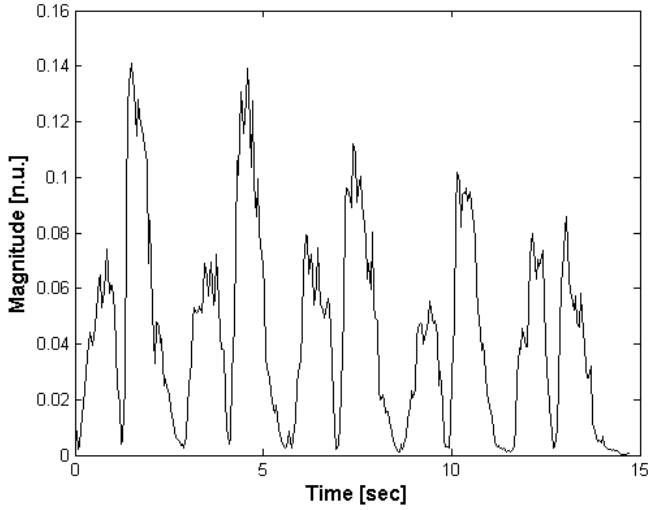


Figure 3-8. Downsampled lung sound signal f(n).

6. As the last pre-processing step, the signal was **de-trended** to remove the global mean and center the signal at amplitude zero to produce *g(n)* (see Figure 3-9).

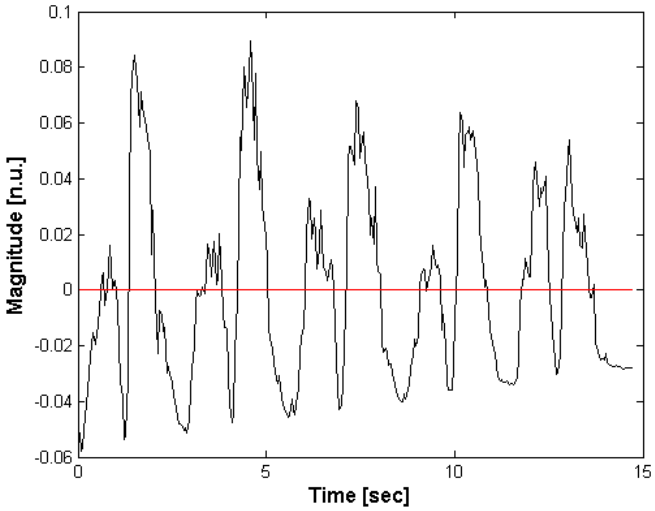


Figure 3-9. De-trended signal g(n) with zero-mean (shown in red for reference).

7. By examining Figure 3-9, one can see that further low-pass filtering is necessary so that the peaks and troughs of each respiratory cycle do not contain multiple turning points. Because adults and pediatric subjects can exhibit a wide range of respiratory rates (from 12-20 cpm for normal adults and from 15-80 cpm for normal pediatric subjects), the Hamming window size for the low pass filter was adapted according to the

characteristics of each recording (Rosdahl and Kowalski 2008). The Hamming window function *hw(n)* was given by Equation 3.4 for $|n| \leq Q$ (Smith 1997).

$$hw(n) = 0.54 + 0.46\cos(n\Pi / Q) \tag{3.4}$$

The window size N_H was calculated using Equation 3.5, where *L* is the length in samples of the widest peak lobe in *g(n)*. The **widest peak lobe** for a signal was the largest length between two zero-crossings over the entire signal (refer to Figure 3-10).

$$N_H = ceil(L/2) + 2 \tag{3.5}$$

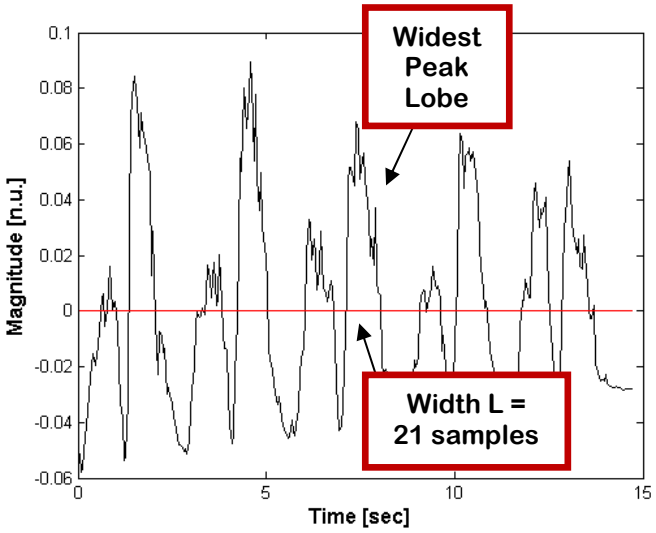


Figure 3-10. Widest peak lobe with length L.

8. After determining the window size N_H, a low pass filter with N_H-point **Hamming window** was then applied to the signal *g(n)* to remove high frequency noise in the envelope and to more clearly distinguish the peaks and troughs. The resulting signal *h(n)* is shown in Figure 3-11.

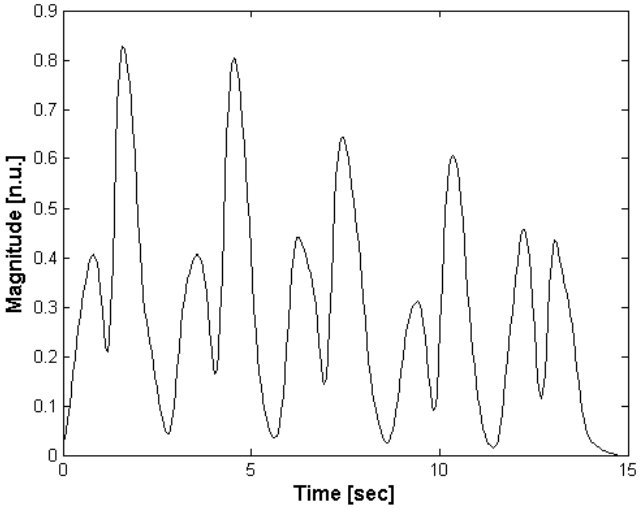


Figure 3-11. Low pass filtered signal h(n) using a N_H -point Hamming window.

9. The signal *h(n)* was then low pass filtered again with the same variable length **Hamming window** (computed as in step 7) to further smooth out the high frequency noise introduced in the signal from downsampling to yield the signal *j(n)* – illustrated in Figure 3-12.

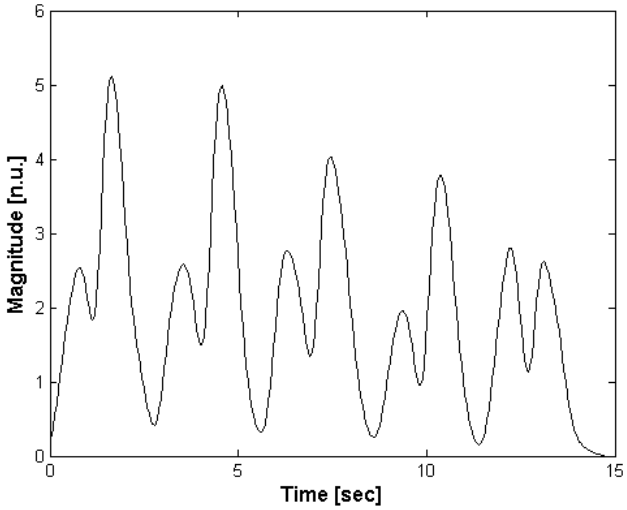


Figure 3-12. Low pass filtering for a second time using a w-point Hamming window to produce j(n).

10. The **single-point derivative** of the filtered signal, *j(n)*, was taken to emphasize slope information (so that a peak could later be identified as a zero crossing). The derived signal *l(n)* is pictured in Figure 3-13.

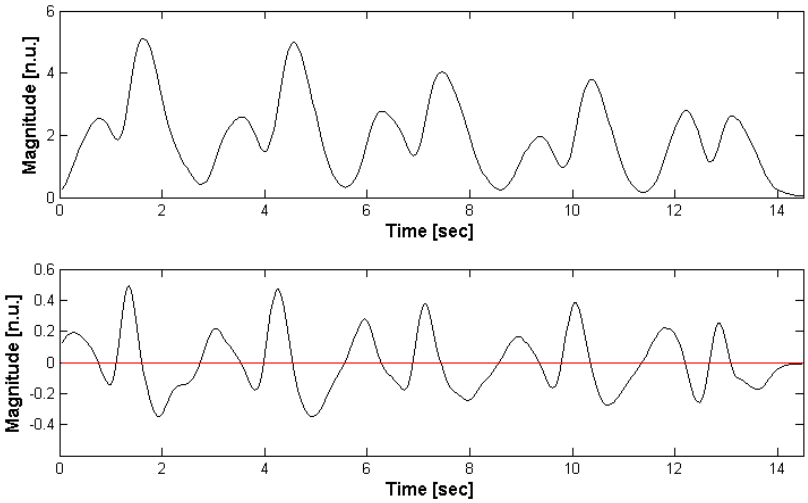


Figure 3-13. The derivative l(n) signal (bottom) of the j(n) signal (top) with the zero-amplitude line as reference (in red).

11. **Valid peaks** are then identified by forming a **binary vector** *m(n)* using Equation 3.6, where λ is a fraction of the max value of the signal *l(n)* that determines the **threshold** that must be passed for consideration as a peak.

$$\begin{matrix} m(i) = 1 & if & l(i) > \lambda * \max(l(n)) \\ 0 & if & otherwise \end{matrix} \tag{3.6}$$

For analysis on this database of recordings, *λ* was set to 0.001. This meant that *m(i)* was equal to one if *l(i)* was ≥ 0.1% of the max value in the entire *l(n)* signal, where the max value represented the max slope in the lung sound envelope signal *j(n)*. If *l(i)* was less than the threshold, then *l(i)* was set to zero (see bottom graph in Figure 3-14). There were eight detected peaks in *m(n)* for this example in Figure 3-14, which was the expected result from observing eight respiratory phases in the raw signal in Figure 3-2.

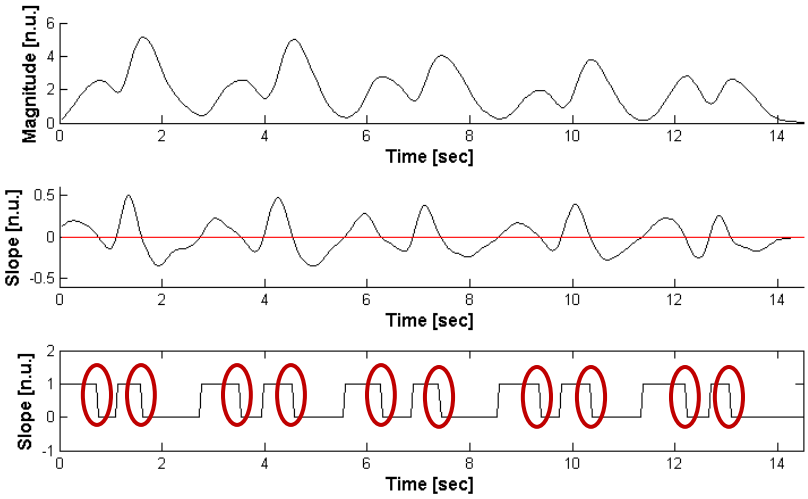


Figure 3-14. The envelope lung sound signal j(n) [top], its derivative l(n) [middle], and the binary vector m(n) [bottom] with the 8 detected valid peaks circled in red in signal m(n).

12. The **respiratory rate** (RR) for the TBC method is given as follows

$$RR = \frac{\left(\frac{N_{valid\_peaks}}{2} \right)}{T} \tag{3.7}$$

With length of signal duration T (in seconds) and N_{valid_peaks} as the number of valid peaks (Yi 2004).

3.2.2 Modified Pan and Tompkins' Method

The Pan and Tompkins' (1985) real-time QRS detection algorithm was explained in Chapter 2.2 for detecting beats in ECG and heart sound signals. The rationale in using this algorithm for respiration rate estimation is based on the observation of high frequency spikes at the points of inspiration and expiration in lung sound signals. Since the Pan and Tompkins' algorithm was developed to detect a series of spikes in the ECG, it is reasonable to assume that the same algorithm could be applied to detect the respiration-related spikes in the audio time series. In the lung sound signal amplitude

plots (see Figure 3-1), each respiration phase appeared as a sudden peak that tapered off to an amplitude of approximately zero, until the next phase abruptly began with a peak and tapered-off in amplitude in a repeating fashion. The signal peaks at inspiration and expiration are thought to be the result of turbulent air flow over large airways in the lungs (Loudon and Murphy 1984). The modified Pan and Tompkins (MPT) method used in this section is very similar to the one explained in Chapter 2.2 and is also based on Clifford's (2002) implementation of a batch QRS detector. The MPT method is outlined in the block diagram in Figure 3-15 and described in five steps below.

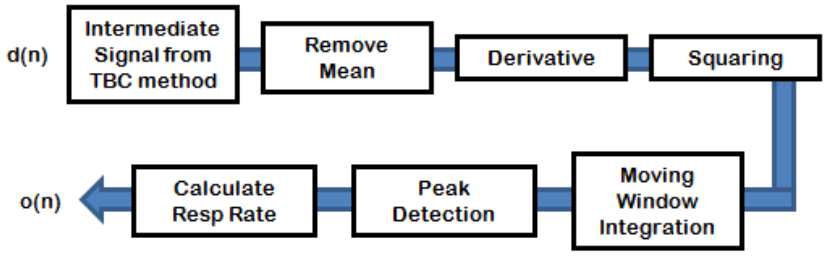
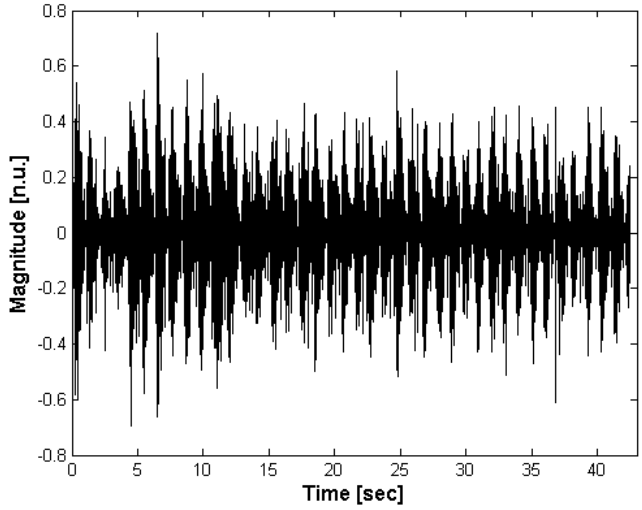
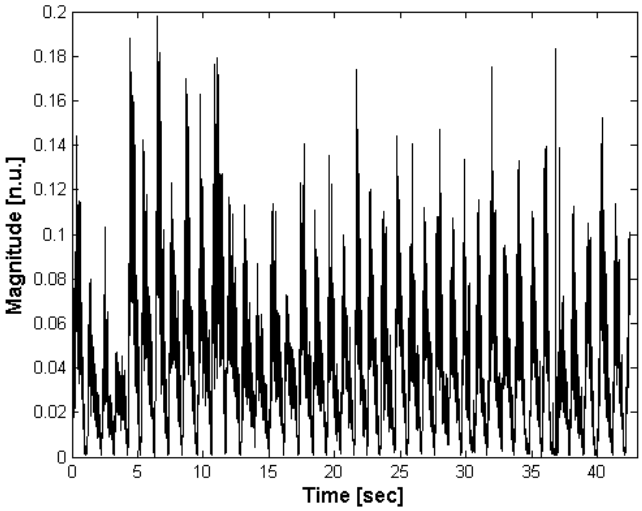


Figure 3-15. Block diagram of MPT method for lung sound respiration rate estimation.

1. The ideal input to this method is a signal with clear peaks at each respiratory phase. As a result, only limited **pre-processing** was performed on the lung sound signal while still maintaining the spikes in the raw signal. The initial steps of the **TBC method** were used to remove noise and improve computation processing time to obtain signal *d(n)* from the block diagram in Figure 3-3 above. Any further filtering would smooth out the signal peaks and lead to poor performance by the peak detector. Figure 3-16a illustrates the raw respiratory sound signal (recording 20 in Table 3-1 with respiratory rate 29 cpm), while Figure 3-16b shows the result of applying the initial stages of the TBC method to the same recording.



(a)



(b)

Figure 3-16. Pre-processed lung sound signal from intermediate stage in TBC method.
a) Raw lung sound signal,
b) Signal *d(n)* according to the block diagram in Figure 3-3.

2. The remaining steps in the MPT method were similar to the QRS detection approach from Chapter 2.2. The **mean was removed** from the signal obtained after step one. Then the **derivative** of the signal was taken to extract slope information. Next, the signal was **squared** to further accentuate the peaks. Figure 3-17 displays the state of the signal after each of these steps.

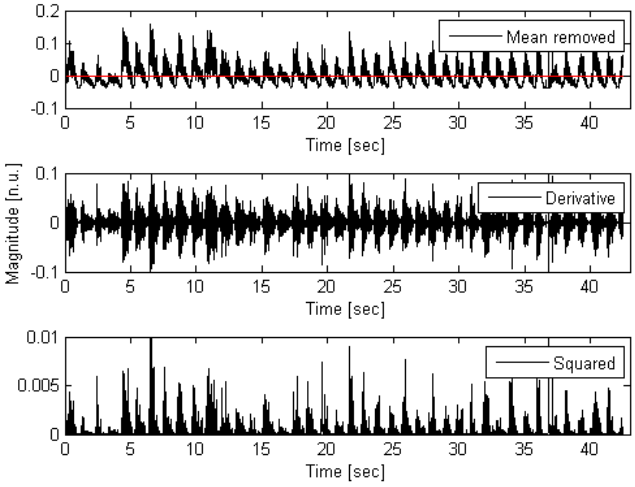


Figure 3-17. Further pre-processing of signal in MPT method to remove mean (top), take the derivative (middle), and square the signal (bottom).

3. The signal was then **integrated** using a moving window. The width of the window is important because too small of an integration window could have caused a single peak in the respiration signal to trigger false peaks with the detector, while too large of an integration window could have missed peaks. For analysis on this database of adult recordings, a moving window size of N=19 samples (see Equation 2.5) was used. Figure 3-18 shows the lung sound signal before (black) and after the integration step (blue).

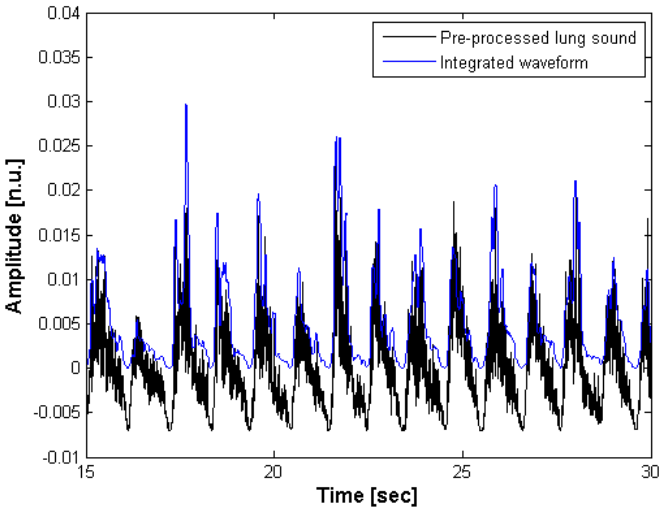


Figure 3-18. Integrated lung sound signal.

4. **Peaks** were then **identified** within large energy bursts in the integrated waveform using an empirical threshold parameter of t=0.21 (see Equation 2.10). Figure 3-19 portrays the peaks and troughs labeled on a breath sound signal. Following peak detection, a list of time intervals (in seconds) between each respiration peak was stored in $T_{peak\_intervals}$.

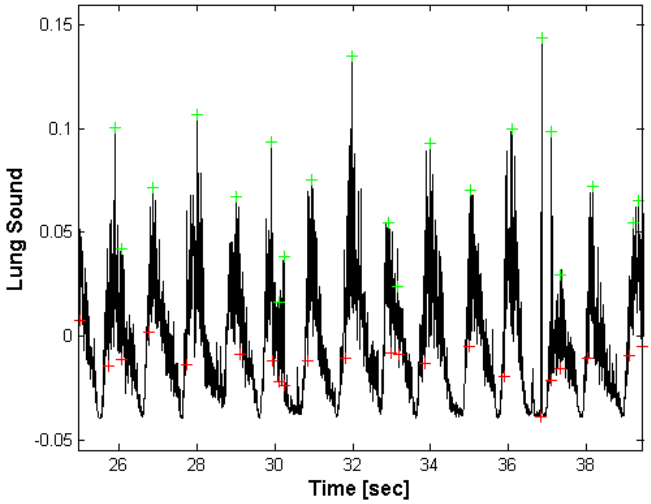


Figure 3-19. Plus signs (+) indicate peaks (green) and troughs (red) on the breath sound.

5. The median time interval (in seconds) between two signal peaks was used to compute **respiration rate** as the number of respiration cycles (two phases per cycle) per minute with Equation 3.8.

$$RR = \frac{\left(\frac{60}{median(T_{\text{peak\_intervals}})} \right)}{2} \tag{3.8}$$

3.2.3 Autoregressive Modeling Methods

The third respiration rate estimation method originated from Fleming and Tarassenko's (2007) work. The two researchers used an autoregressive (AR) model approach involving spectral analysis to extract respiration rate from a PPG waveform – a standard vital sign collected in hospitals using a pulse oximeter. Their algorithm was tested on 14 five-minute sections of data from the records of seven adults in the Physiobank MIMIC database (Goldberger 2000). Each record contained simultaneously captured PPG and respiratory waveform data, the latter of which was used as the reference signal and recorded with a nasal thermistor (Fleming and Tarassenko 2007). The AR model approach was able to accurately estimate respiratory rate from the PPG signal with a mean error of 0.04 cpm.

The AR model attempts to represent the signal as a linear prediction equation *y(n)*, which is a sum of *p* previous values (where *p* is the model order), weighted according to the AR model coefficients q_i, and combined with an estimation error *err(n)* – see Equation 3.9 (Acero 1993, Fleming and Tarassenko 2007).

$$y(n) = -\sum_{i=1}^{p} q_i \, y(n-i) + err(n) \tag{3.9}$$

The z-transform of the AR model can be expressed as an all-pole transfer function where C_{error} is the square root of the estimation error – refer to Equation 3.10 (Acero 1993). Each root z_i in the denominator signifies a pole in $H(z)$, which indicates a peak in the power spectrum of the respiration signal (Fleming and Tarassenko 2007). The frequency f of the peak is determined according to the angle Θ at which the pole occurs using Equation 3.11, where t is the sampling time interval.

$$H(z) = \frac{C_{error}}{1 + \sum_{i=1}^{p} q_i z^{=i}} \tag{3.10}$$

$$\Theta = 2\Pi ft \tag{3.11}$$

Since a breathing trace manifests as a naturally quasi-sinusoidal process, there is a dominant frequency at which an individual will inhale and exhale. This frequency value can be converted to a respiratory rate in cpm. AR modeling allows identification of the dominant frequency within the breathing waveform with relatively few samples as described below.

Fleming and Tarassenko's (2007) original algorithm started by **band pass filtering** the raw PPG signal between 0.1Hz and 40Hz to reduce noise and interference from heartbeats. The filter was implemented with 490 coefficients using a Kaiser window function. For additional pre-processing, the resulting signal was **downsampled** to 1 Hz and **detrended** to increase low frequency data resolution (corresponding to respiration rate information) and increase model stability. The data was then fitted to an **AR model** as described in Equations 3.9 – 3.11, where they determined that an order 11 model would yield the lowest error for respiration rate extraction from a PPG signal.

An acceptable **range of breathing rates** for adults was identified to be between 4.8 to 42 cpm (Fleming and Tarassenko 2007). This range was translated into an acceptable range of angles for the poles in the z-plane from the *H(z)* transfer function. In order to identify the pole corresponding to the respiration rate of the patient, the group of candidate poles (which signified the resonant frequencies in the original respiration signal) was narrowed down according to those with angles that fell within the range of physiological plausibility. Higher magnitude poles in the transfer function correspond to higher magnitude peaks in the power spectrum, so only the most dominant frequencies with highest power spectrum magnitudes were desired (as likely representing the respiratory rate of the signal). Hence, the poles with magnitude greater than or equal to a **threshold** (set at 95% of the highest magnitude pole among those within the area of interest) were kept for consideration. Of the remaining poles, the pole with the **smallest angle** (or lowest frequency) was finally selected as the respiration rate of the PPG waveform.

A modified implementation of Fleming and Tarassenko's AR-based approach was developed in Nemati et al.'s (2010) Kalman Filter fusion framework software for robust fusion of respiratory rate estimations from data collected through the ECG, PPG, and impedance pneumogram. Although the respiration rate estimation algorithm used in AR model method for this thesis used the source code in the work of Nemati et al. with relatively few modifications, the data fusion approach presented here was not based upon their Kalman Filter-fusion framework. Below is a description of the AR model method 1 (ARM1) used for extracting respiratory rate from acoustic lung sounds with a

sliding 10-second window and 5-second overlap. The block diagram for ARM1 is displayed in Figure 3-24.

1. Nemati et al.'s method for AR model processing was used with an input of the **envelope of the lung sound signal** just after two rounds of low pass filtering in the TBC method (Section 3.2.1), otherwise known as signal *j(n)* in Figure 3-3. The signal *j(n)* was sampled at 8 Hz (from an original raw signal sampled at 8000 Hz). Figure 3-12 illustrates an example of a respiratory sound signal envelope that would be used.

2. The data are **detrended** (see Figure 3-20) to have a zero global mean to avoid the clustering of poles at low frequencies near DC.

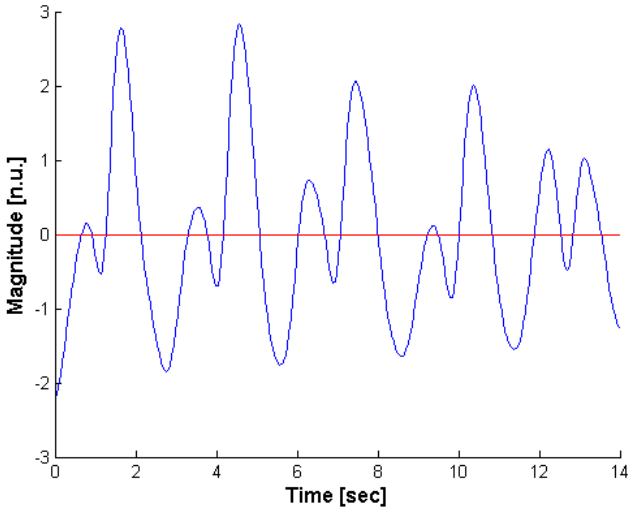


Figure 3-20. The input ARM1 signal detrended.

3. The resulting signal is fit to an **18th-order AR model**, so that the 18 most dominant frequencies can be identified. These poles are plotted on the pole-zero diagram in the z-plane shown in Figure 3-21.

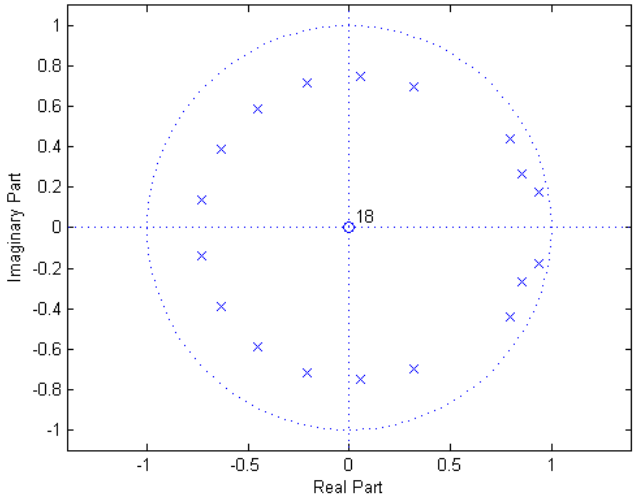


Figure 3-21. Poles for 18th-order AR model fitted on lung sound recording waveform.

4. The acceptable **breathing range** used in this thesis is 2 cpm – 35 cpm for adults and 2 cpm – 80 cpm for pediatric subjects (< 12 years old) – see Section 3.3.2 for further elaboration.

5. Given the 18 poles and their angular information, the frequencies of each pole were calculated using Equation 3.11. Only the poles representing **frequencies** within the **range of physiological plausibility** (from 3º - 26º corresponding to 2 - 35 cpm) were kept for consideration as in Figure 3-22.

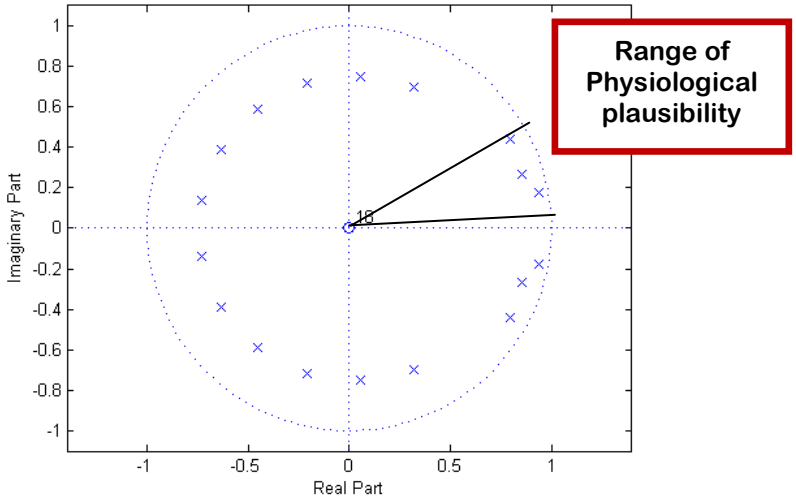


Figure 3-22. Exclude poles outside acceptable breathing range.

6. The largest magnitude pole within the range of interest was assumed to correspond to the **respiratory frequency**. In the example, the pole within the range of interest with the lowest angle is selected as the respiratory rate estimate (marked in Figure 3-23).

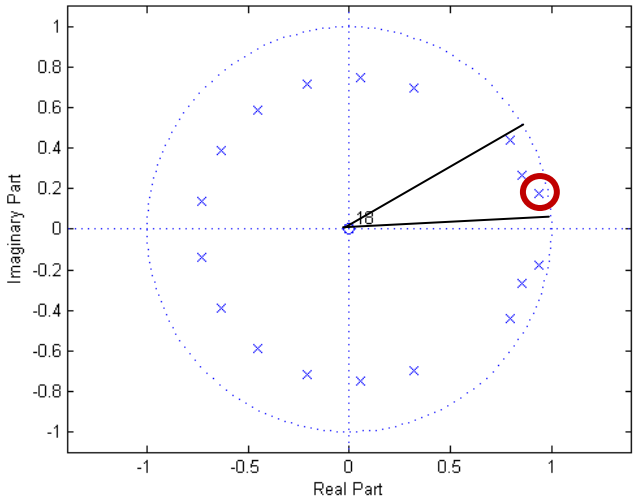


Figure 3-23. Pole representing respiratory rate is circled in red.

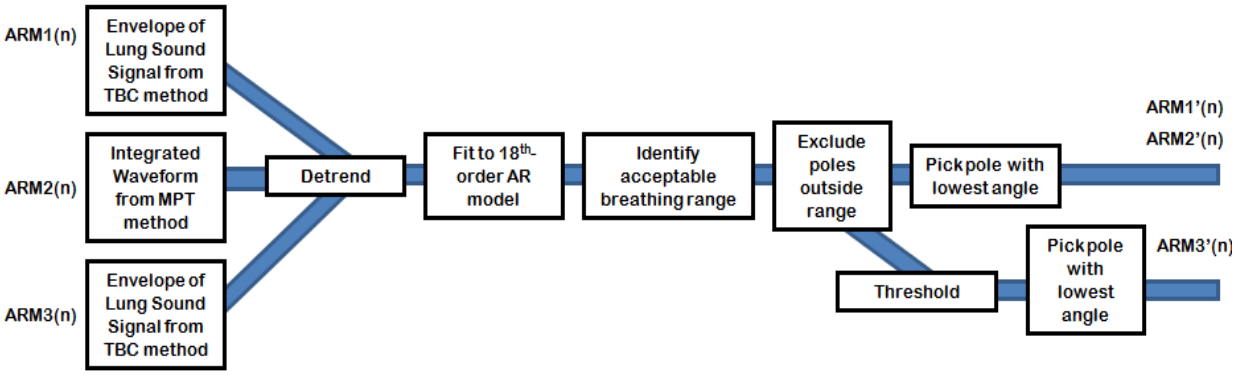


Figure 3-24. Block diagram of ARM methods one, two, and three.

Three variations of the AR model approach were implemented and tested in an attempt to improve algorithm performance (refer to Figure 3-24). ARM1, as described in steps 1 to 6 above, attempted to fit an AR model to the smooth envelope of the lung signal from an intermediate step *j(n)* in the TBC method. The Autoregressive Model 2 (ARM2) method employed the same techniques as ARM1, except ARM2 fit an AR model to the intermediate lung sound waveform from the MPT method immediately after

the moving integration window stage (blue line marked in Figure 3-18). In the Autoregressive Model 3 (ARM3) method, the envelope of the lung sound signal from the TBC method was used (as in ARM1). However, only the poles with magnitude greater than or equal to 95% of the highest magnitude pole in the range of physiological plausibility were kept as candidates. Then the respiratory rate frequency was selected as the pole with the lowest angle among the remaining poles.

3.2.4 Product of Experts

Since a voting scheme between several algorithms often out-performs any single algorithm (Moody et al. 2006, Moody and Lehman 2009), another method that was tested was a simple median of the three algorithms described in sections 3.2.1 to 3.2.3 (TBC, MPT, and ARM1). This scheme is referred to as the MEDIAN method in this analysis. The rationale here was that if only one algorithm performed badly for a particular lung sound recording, the poor result would be discarded by the median method because the other two methods would hopefully yield a closer estimate to the actual respiratory rate.

3.2.5 Signal Purity

To ensure that the respiratory rate detector does not attempt to process a very noisy signal and report a misleading result, it is important to first determine whether the signal is of sufficient quality. Since a clean respiratory recording should be an approximately sinusoidal waveform, its signal quality can be quantified using its spectral width, or *purity* (Nemati et al. 2010). The purity value of a waveform is equal to one for a perfectly sinusoidal signal of exactly one frequency and degrades to zero for a noisy

signal. In a noisy signal, the bandwidth of the signal increases while non-stationarity, distortion and outside noise become more prevalent.

A signal's purity index is computed using Hjorth descriptors, which summarize the spectral characteristics of a waveform through the total signal power, spectral width, and dominant frequency (Hurtig-Wennlof et al. 2009, Nemati et al. 2010). These Hjorth parameters are computed from nth order spectral moments determined by Equation 3.12 (Nemati et al. 2010) where S_x is the spectral power for a signal. For purity value computation in Equation 3.13, the 0th, 2nd, and 4th spectral moments were used.

$$\overline{\omega_n} = \int_{-\Pi}^{\Pi} \omega^n S_x(e^{jw}) d\omega \tag{3.12}$$

$$purity(k) = \frac{\overline{\omega_2^2}(k)}{\overline{\omega_0}(k)\overline{\omega_4}(k)} \tag{3.13}$$

Nemati et al. (2010) also used this signal quality index measure in their software for respiratory rate estimation described in Section 3.2.3. Their framework automatically estimated respiratory rate from each of three physiological sources (ECG, IP, and PPG) and was able to fuse them into a single robust estimation by weighting the estimation from each source using a signal quality index. Their implementation for signal quality was used as part of this lung sound analysis research (see block diagram 3-25).

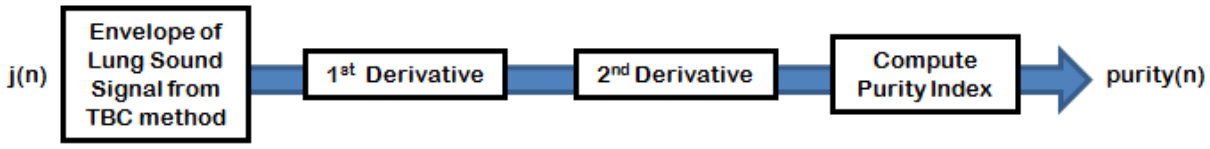


Figure 3-25. Block diagram of signal purity calculation.

1. Since respiration should be a slow-moving sinusoid, the purity index was evaluated for the **envelope of the acoustic lung signal** - *j(n)* from the TBC method - instead of the raw signal with high frequency noise. To compute the nth order spectrums (for n=0,

2, and 4), the **first and second derivative** of the envelope signal was taken. Figure 3-26 shows an example of a signal and its two derivatives, corresponding to recording 20 in Table 3-1.

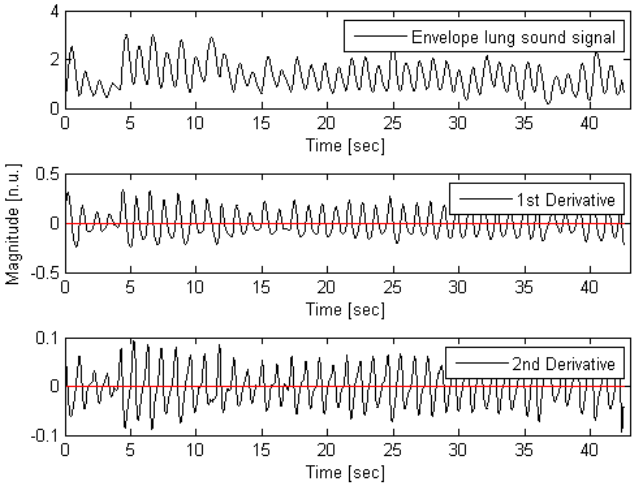


Figure 3-26. Envelope of lung sound signal, its first derivative, and its second derivative.

2. These three signals were then used to **compute the purity value** in Equation 3.13 while solving for w_0, w_2, and w_4 using Equation 3.12. When applied to the example signal represented in Figure 3-26, the purity index was calculated to be 0.10.

From assigning a purity score to each signal in the lung sound dataset, a threshold could be developed such that a signal with purity lower than that threshold must be discarded for too much noise and outside interference distortion. This threshold was determined empirically and is discussed in the next section.

3.3 Analysis of gold standard database

The TBC, MPT, ARM, and MEDIAN methods were evaluated across all 74 recordings from the lung sound database. The performance of each method was

reported in terms of percentage error between the algorithm's estimated respiration rates for the recordings and the manually computed respiration rates. Since the three methods were originally developed for purposes other than lung sounds (such as QRS complex detection in ECG waveforms), changes were introduced to tailor the methods for respiration rate extraction. The modifications are described in Sections 3.3.1 and 3.3.2.

3.3.1 Pre-processing of the respiration signal

Since the lung sound audio data had to be manually recorded with an electronic stethoscope, there was sometimes a 'dead-space' interval before and after the lung sound content in the resultant WAV files (see recording 53 from subject 10 in Figure 3-27). To remove the intervals that lacked lung sound information, a MATLAB script was written to take the raw signal and truncate the sections with low amplitude values at the start and end of the signal.

This "Truncate Signal" stage consisted of scanning forwards through the signal from the start until lung signal content was detected. Lung sound content was detected by computing the variance of a moving window (of 100 samples in width) and comparing it against the variance of the entire recording. If the variance in the window was greater than 1% of the variance of the overall signal, then actual lung sound content was assumed to be present. This was repeated backwards from the end of the signal. The result was a new shortened signal with the lung sound content only (see Figure 3-28).

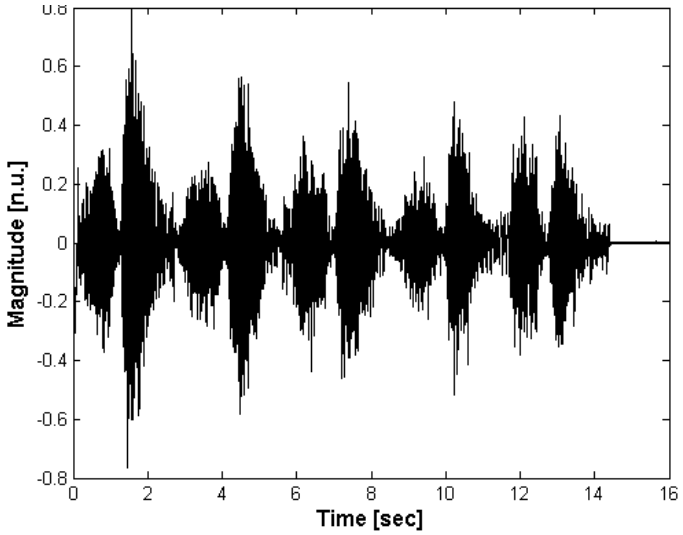


Figure 3-27. Before the Truncate Signal Stage.
Original WAV file with no lung sound content in last second of recording.

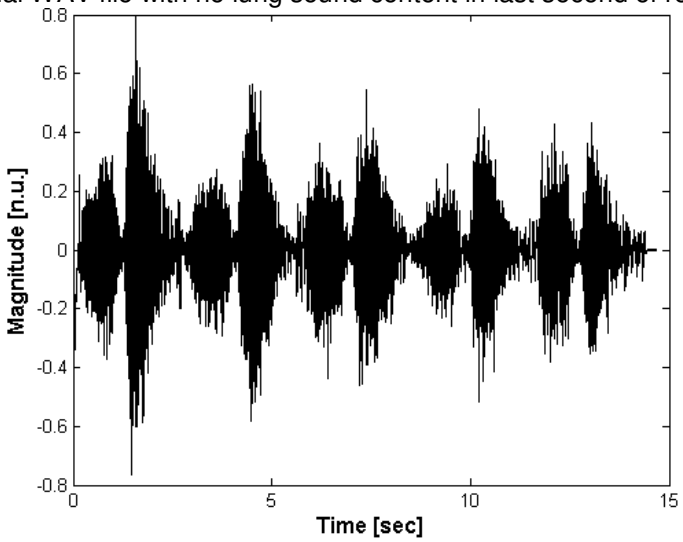


Figure 3-28. After the Truncate Signal stage.
Same WAV file as in Figure 3-27 with dead space removed

3.3.2 Upper and Lower Limits for Respiratory Rate Estimates

From preliminary analysis on the WAV files, the algorithms would occasionally produce estimates that were physiologically improbable. It was therefore decided that a range of acceptable respiratory rates would be integrated into the algorithm to provide an intelligent check on the validity of the estimate.

Ideally the acceptable respiratory ranges for the lung sound database analysis should include possible respiratory rates for both *normal* and *abnormal* patients

because the software was designed for use on both groups. Since the literature did not provide standard ranges, these ones used in this research were developed based on the acceptable ranges for *normal and healthy* individuals by adding a delta of 15 cpm lower than the 1st percentile and above the 99th percentile respiratory rate values for healthy individuals (Fleming et al. 2010, Walter et al. 2004). Ranges are highly dependent on age, so for adults, the valid respiratory rate range was set from 2 to 35 cpm. Pediatric subjects (≤ 12 years old) were not present in this database, but their range was set from 2 to 80 cpm for future lung sound recordings from this age group. A delta of 15 cpm was determined from observation of respiratory rates of subjects in the database and accounting for individuals who would have an even lower or higher respiratory rate than those in the database. However, 2 cpm was used as the absolute lower limit, which would mean 1 respiratory cycle per 30 seconds. This rate is extremely slow and would be noticeable by the attending healthcare worker by watching the patient breathe (without aid from the phone software). For reference on the acceptable respiratory ranges for all ages, view Sections 1.2.2 and 1.3.2 of this thesis.

If an algorithm produced a value above or below the range, the maximum (or minimum) acceptable value would be returned instead. Exact values outside these ranges are unimportant, since knowing that the person is likely exhibiting a respiratory rate at the very extremes of the range is sufficient to indicate abnormality and would necessitate medical intervention.

3.4 Results

Table 3-3 presents the final results for the performance of each algorithm (TBC, MPT, ARM1, ARM2, ARM3, and MEDIAN) on respiratory sounds from the standard database. Results are reported in percentage errors because absolute error values (in cpm) overweight high respiratory rates. By taking the error as a percentage of the actual respiratory rate, the error was normalized across recordings of low and high respiratory rate. The last column in the table also expresses the purity level of the signal. Table 3-4 summarizes the performance of each method on respiratory rate estimation for the recordings.

ID	ACTUAL CPM	TBC	MPT	ARM1	ARM2	ARM3	MEDIAN	PURITY
1	15	0.00	0.71	0.79	0.96	0.79	0.71	0.12
2	17	-0.06	1.06	0.73	0.78	0.73	0.73	0.07
3	15	-0.07	1.33	0.95	0.97	0.95	0.95	0.07
4	15	-0.07	1.33	0.65	1.08	0.65	0.65	0.06
5	15	0.00	1.33	0.97	1.03	0.97	0.97	0.05
6	14	0.36	1.33	1.09	1.32	1.09	0.72	0.16
7	15	0.53	-0.31	1.24	1.16	1.24	0.53	0.04
8	11	-0.09	2.18	1.01	1.80	1.01	1.01	0.08
9	12	0.00	-0.56	1.21	1.65	1.21	0.00	0.07
10	14	0.21	-0.84	0.83	1.30	0.83	0.21	0.01
11	23	-0.30	0.16	0.16	0.33	0.16	0.16	0.05
12	23	-0.04	0.52	-0.91	0.35	-0.91	-0.04	0.11
13	25	-0.04	0.40	-0.92	0.25	-0.92	-0.04	0.06
14	25	-0.04	-0.74	-0.92	0.22	-0.92	-0.74	0.05
15	25	0.08	0.40	-0.92	0.26	-0.92	0.08	0.04
16	25	-0.04	0.34	-0.92	0.20	-0.92	-0.04	0.11
17	26	-0.08	0.35	-0.92	0.19	-0.92	-0.08	0.07
18	27	-0.04	0.30	-0.93	0.18	-0.93	-0.04	0.06
19	26	-0.04	0.27	-0.92	0.09	-0.92	-0.04	0.1
20	29	-0.03	0.21	-0.93	-0.03	-0.93	-0.03	0.1
21	25	-0.64	-0.38	-0.18	0.16	-0.18	-0.38	0.05
22	16	0.25	-0.30	-0.88	0.90	-0.88	-0.30	0.08
23	17	0.24	-0.85	-0.88	0.78	-0.88	-0.85	0.04
24	18	0.78	0.39	0.62	0.78	0.62	0.62	0.02
25	12	0.00	1.92	1.28	1.43	1.28	1.28	0.14
26	17	-0.12	-0.80	0.63	0.89	0.63	-0.12	0.02
27	17	-0.18	1.06	0.46	0.83	0.46	0.46	0.03
28	14	0.79	1.50	-0.86	0.99	-0.86	0.79	0.02
29	18	-0.17	0.94	0.56	0.67	0.56	0.56	0.03
30	14	0.64	1.50	1.05	1.30	1.05	1.05	0.03
31	15	0.00	1.50	0.97	1.00	0.97	0.49	0.07
32	14	0.36	1.50	1.14	1.15	1.14	1.14	0.06
33	12	0.00	0.74	1.25	1.19	1.25	0.74	0.09

34	13	0.00	1.69	1.22	1.21	1.22	1.22	0.11
35	13	0.08	1.69	1.22	1.20	1.22	1.22	0.12
36	15	0.00	0.41	1.31	0.97	1.31	0.41	0.15
37	12	1.92	-0.59	-0.83	1.63	-0.83	-0.59	0.05
38	15	-0.27	-0.19	0.78	0.96	0.78	-0.19	0.19
39	16	0.94	-0.81	-0.88	0.99	-0.88	-0.81	0.03
40	13	1.46	1.69	-0.85	1.33	-0.85	1.46	0.02
41	12	0.08	0.75	1.44	1.52	1.44	0.75	0.19
42	13	0.54	-0.43	1.20	1.48	1.20	0.54	0.06
43	16	0.31	1.19	0.98	0.85	0.98	0.98	0.1
44	15	0.07	1.33	1.02	1.09	1.02	1.02	0.08
45	14	1.07	1.50	-0.86	1.23	-0.86	1.07	0.06
46	16	0.81	0.95	0.05	0.74	0.05	0.81	0.01
47	14	0.00	0.95	1.22	1.19	1.22	0.61	0.06
48	14	0.57	1.50	0.99	1.04	0.99	0.99	0.06
49	15	0.07	1.33	-0.87	0.91	-0.87	0.07	0.11
50	15	-0.07	-0.84	1.21	1.10	1.21	-0.07	0.13
51	19	-0.05	0.84	-0.89	0.71	-0.89	-0.05	0.1
52	18	0.17	0.90	-0.89	0.68	-0.89	0.17	0.06
53	15	0.33	1.33	-0.87	0.91	-0.87	0.33	0.13
54	17	0.00	1.06	1.02	0.83	1.02	1.02	0.18
55	21	0.67	-0.37	-0.90	0.27	-0.90	-0.37	0.17
56	22	0.32	-0.37	-0.91	0.30	-0.91	-0.30	0.06
57	25	-0.44	0.15	0.28	0.31	0.28	0.15	0.05
58	19	0.58	-0.67	-0.11	0.64	-0.11	-0.11	0.06
59	23	0.09	0.52	-0.91	0.39	-0.91	0.09	0.04
60	21	0.48	0.67	0.61	0.38	0.61	0.61	0.03
61	30	-0.33	0.17	0.14	-0.06	0.14	0.14	0.08
62	21	-0.14	-0.35	0.35	0.50	0.35	-0.14	0.07
63	17	0.00	1.06	-0.04	0.72	-0.04	0.00	0.08
64	14	0.07	1.41	1.16	1.18	1.16	1.16	0.09
65	6	0.00	3.05	1.89	4.32	1.89	1.89	0.15
66	8	0.00	3.38	1.65	2.83	1.65	1.65	0.11
67	8	0.00	3.38	1.49	2.83	1.49	1.49	0.21
68	13	-0.08	-0.05	1.05	1.25	1.05	-0.05	0.23
69	15	-0.07	-0.25	1.01	0.89	1.01	-0.07	0.19
70	17	-0.18	0.44	0.70	0.50	0.70	0.44	0.11
71	15	-0.07	1.33	1.03	0.89	1.03	1.03	0.18
72	12	0.00	1.92	1.24	1.45	1.24	1.24	0.21
73	18	0.06	-0.09	-0.89	0.69	-0.89	-0.09	0.13
74	18	0.11	0.50	-0.89	0.76	-0.89	0.11	0.06

Table 3-3. Respiratory rate estimation results on respiratory sounds from standard database.

Method	Recordings Used [%]	Average Absolute Error [%]	Root Mean Squared Error [%]	Root Mean Squared Error [cpm]
TBC	100	25.33	4.34	6.70
MPT	100	94.88	11.80	15.34
ARM1	100	89.95	9.64	15.34
ARM2	100	94.53	11.53	14.21
ARM3	100	89.95	9.64	15.34
MEDIAN	100	56.89	7.36	9.94

Table 3-4. Summary of results for respiratory rate estimation

The results demonstrate that the TBC methods performed the most accurately on respiratory sounds in the database with a root mean squared percentage error (RMSPE) of 4.34% and root mean squared error (RMSE) of 6.70 cpm. These error values can be compared with the results from the MEDIAN method, which was the next best method with RMSPE 7.36% and RMSE 9.94 cpm. The next best methods were the ARM1 and ARM3 methods (with RMSPE 9.64% and RMSE 15.34), followed by the ARM2 (RMSPE 11.53% AND RMSE 14.21 cpm) and MPT method (RMSPE 11.80% and RMSE 15.34 cpm).

Linear regression was performed on the absolute percentage error rates for each method and graphed on the plots in Figure 3-29. Each plot shows the least squares linear fit (in blue). Interestingly, Figure 3-29 shows that the best fit line for the TBC method (the best estimator) is approximately flat. This observation suggests that the TBC method would fortunately not need to be adjusted for adult patients across the entire range of respiratory rates. For the MPT, ARM1, and MEDIAN methods, there were high correlation values (R^2=0.465, R^2=0.302, and R^2=0.432 respectively) for a trend of decreased absolute percentage error rates with increased actual respiration rate from the subjects.

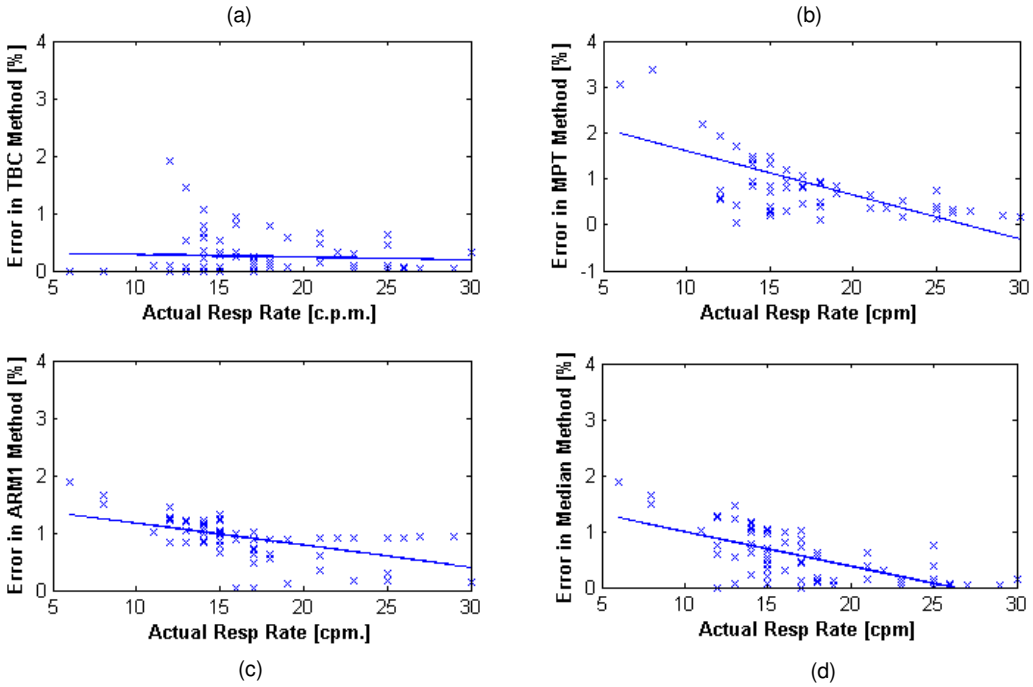


Figure 3-29. Actual Resp Rate (cpm) versus Absolute % Error.
a) TBC method, best fit line y=0.335-0.005x with R^2=0.005
b) MPT method, best fit line y=2.579-0.96x with R^2=0.465
c) ARM2 method, best fit line y=-1.549-0.038x with R^2=0.302
c) MEDIAN method, best fit line y=1.614-0.061x with R^2=0.432

Purity level is another important factor to consider because if a signal is primarily composed of noise, then running even the most advanced signal processing algorithms on it will not produce a reasonable result. The purity measure is defined as a measure of how sinusoidal a signal is. On the database of recordings, the purity level ranged from 0.01 to 0.23 with an average of 0.09. It should be noted that these purity levels are relatively low and correspond to poor expected respiration estimates (Nemati et al. 2010).

Since these recordings were performed electronically with a high-quality commercial microphone, it is expected that mobile phone recordings will have poorer signal quality. To warn healthcare workers about the quality of the signal being recorded, a system (such as the phone-based one considered in this work) will need to

compute the purity level automatically. If the recording is deemed to be poor quality, then the healthcare worker can immediately repeat the recording before transmitting it to a central server for medical expert diagnosis. In order to determine a purity threshold for recordings, the relationship between purity and error rates from each algorithm was plotted, and a linear regression was performed to approximate any trend (refer to Figure 3-30).

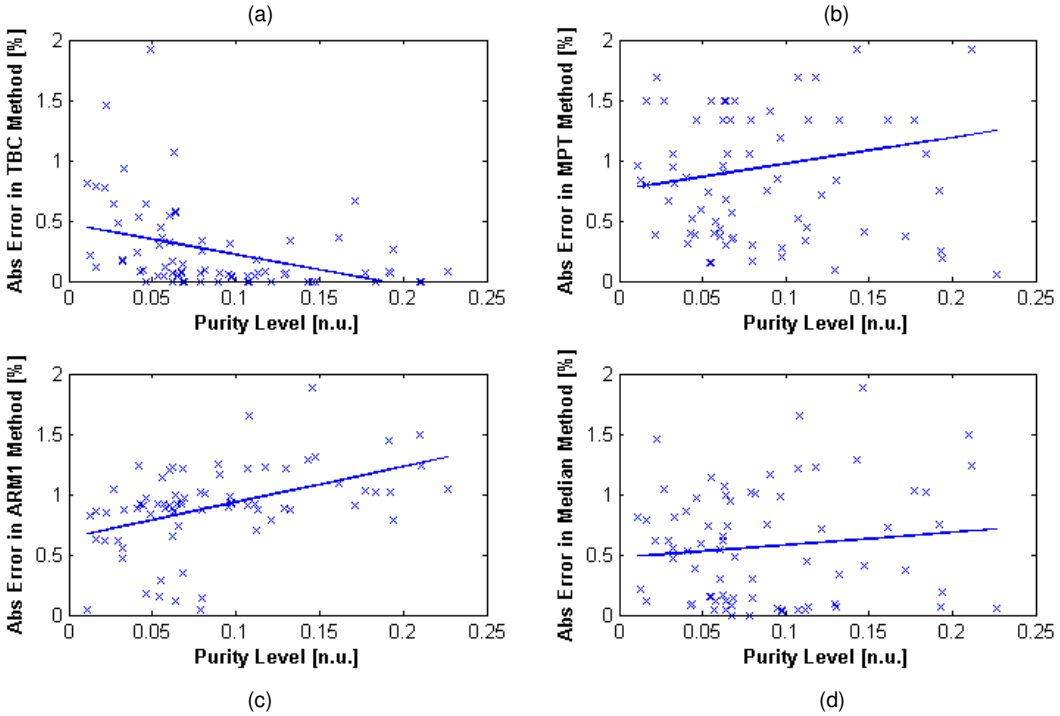
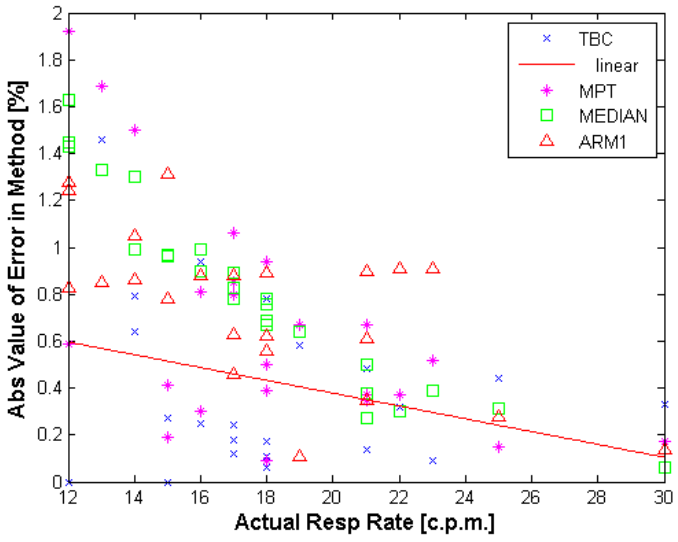


Figure 3-30. Purity level of recordings versus Abs % Error for each method.
a) TBC method, best fit line y=0.476-2.54x with R^2=0.146
b) MPT method, best fit line y=0.758+2.184x with R^2=0.027
c) ARM1 method, best fit line y=0.642+2.939x with R^2=0.202
d) Median method, best fit line y=0.478+1.042x with R^2=0.014

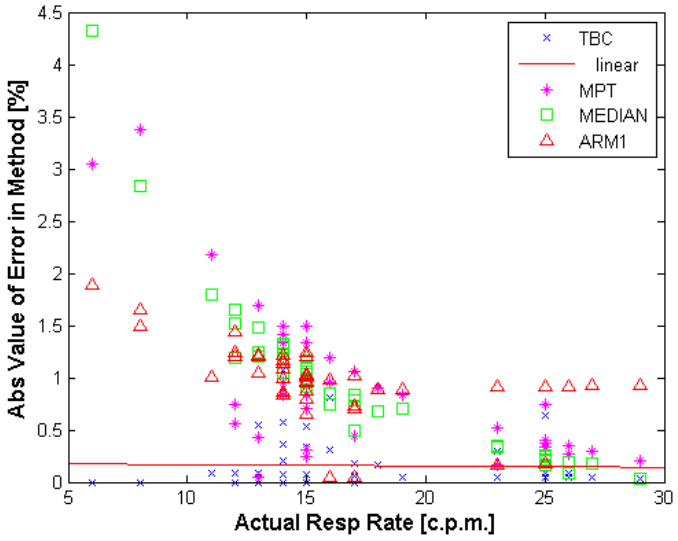
There was not a strong correlation between purity level and estimation accuracy of the MPT and MEDIAN methods, since the data points occurred over a large spread. However, Figure 3-30a shows a promising relationship between purity level and the TBC method estimation error such that signals with higher purity seem to have lower error rates. With the TBC approach, it may be possible to suggest a threshold where

lung sound recordings with purity level ≥ 0.07 would be of sufficient quality to yield a low estimation error. Conversely, the ARM1 method appears to yield estimates with low error for low purity signals. Hence for signals below the purity threshold (i.e. with purity from 0.01 to 0.07), it may be best to use the ARM1 method. Caution should be taken over interpreting these results because any threshold on minimum purity level required would certainly change as more data was collected. Moreover, data collected on mobile phones would likely have different noise profiles and a different purity value may be required. Using the Sana mobile phone capture framework, data would be acquired and labeled as either sufficient quality for diagnosis or not and aid in developing more fine-tuned algorithms to determine signal quality.

To examine the influence of gender on algorithm performance, the data were separated into recordings from females versus males. The relationship between actual respiratory rate and absolute percentage error for all estimators is shown in Figure 3-31. A linear fit is also displayed (in red) on the plots for the TBC method (the best estimator). In recordings with female subjects (Figure 3-31a), absolute percentage estimation error decreased with increased respiratory rate, while there was no significant relationship in the data for male subjects (in Figure 3-31b).



(a)

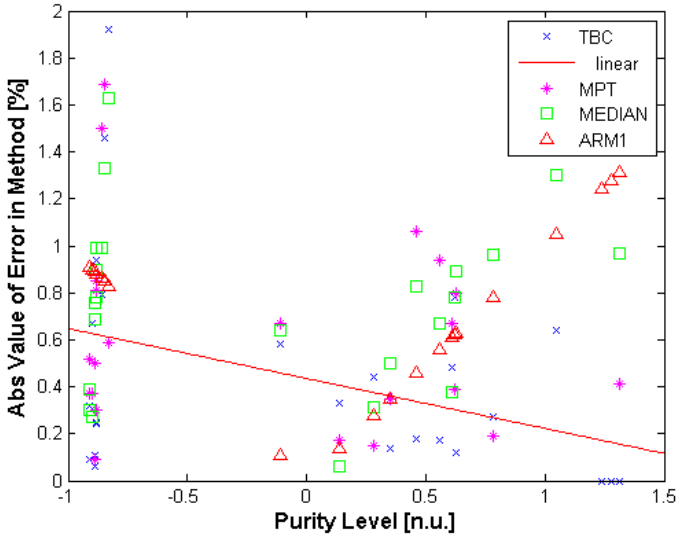


(b)

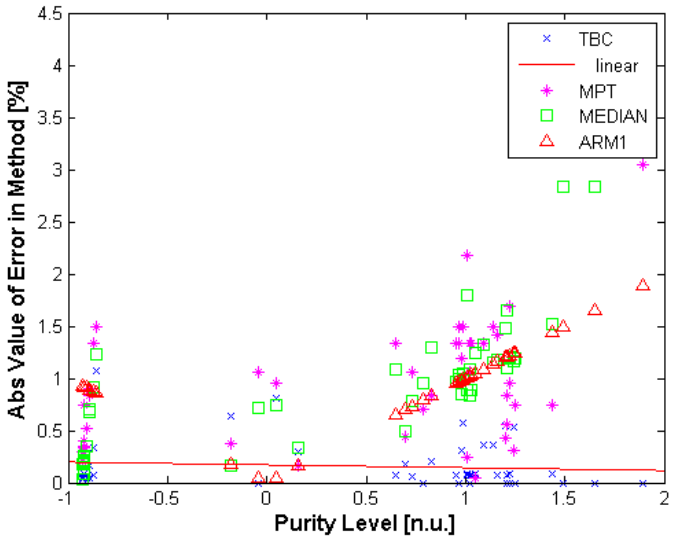
Figure 3-31. Relationship between actual respiratory rate and abs % error according to gender.
a) Female subjects, linear fit for best estimator (TBC method) is y=-0.027*x+0.920 with R^2=0.191.
b) Male subjects, linear fit for best estimator (ARM1 method) is y=-0.001*x+0.180 with R^2=0.084.

The relationship between purity level and percentage error for all the methods was also plotted separately for males and females in Figure 3-32a and b. For the best fit line based on the TBC method results, there was a slight trend in female subjects for decreased error rates with increased purity levels. However, the clustering of data points on the bottom half of the plot in Figure 3-32b implies that the respiration rate

estimates for male recordings produced lower error rates across the range of purity values (and across the different algorithms) compared to female recordings (with a large spread of error values seen in Figure 3-32a).



(a)



(b)

Figure 3-32. Relationship between purity level and abs % error according to gender.
a) Female subjects, linear fit for best estimator (TBC method) is y=-0.21*x+0.440 with R^2=0.194.
b) Male subjects, linear fit for best estimator (ARM1 method) is y=-0.027*x+0.170 with R^2=3.14e-4.

3.5 Discussion

The method that performed the best on respiratory rate estimation on adult recordings in the database was the TBC method. This may be due to the fact that the TBC method was designed specifically for respiratory sounds, while the other methods were adapted from algorithms intended for PPG or ECG waveforms. Since a portion of the data was used by Yi (2004) in the development of her TBC method, it is possible that her algorithm was fitted to perform well on this dataset. However, she used 300 respiratory phases, which was 24% of the 1259 respiratory phases used in this analysis. She may have not chosen to use the other recordings because they were gathered for initial testing purposes (since they had varying duration and noise levels). In this analysis however, all the recordings were useful in order to develop algorithms robust enough for estimation in unpredictable and noisy environments. It is also unknown (and potentially unlikely) that these recordings were captured in a controlled clinical setting (i.e. a doctor's office) by a trained medical professional. For future algorithm development by researchers, a database of normal and abnormal lung sounds should be captured in a standardized clinical setting from patients and made publicly available. This environment may also have other monitoring equipment available to record other vital signals simultaneously, such as respiratory flow from a pneumotachograph to serve as a gold standard in computing the patient's actual respiratory rate. Despite the limitations of the lung sound database used in this research, the results provide a foundation on which to further develop the automated algorithms with larger datasets collected from the field (see Chapter 5).

The MEDIAN method was the second best approach in this analysis and seems promising for application to other lung sound databases, as it would avoid the overfitting issues in this database from Yi's work. The strength of the MEDIAN method estimator was an intuitive observation because no single method was entirely noise insensitive and the MEDIAN method was able to exploit the differing frequency responses of each algorithm. The MEDIAN method could properly balance out the error from each of the three methods to achieve a low estimation error rate for patients.

All of the ARM methods exhibited weaker performances compared to the TBC and MEDIAN methods. One possible explanation for this behavior was the existence of periods (up to three seconds long) without lung sound content in the beginning and end of the recordings due to timing issues in the recording process (i.e. not breathing immediately when the recorder started). These segments did not contain complete silence, but rather noise from the environment and equipment, so they were not completely removed by the preprocessing step in Section 3.3.1. As a result, the database recordings may not have had sinusoidal waveforms (an ideal fit for the AR model) if the lung sound content only occurred in the middle of the recordings (amidst other continuous noise artifacts). The MPT method performed the poorest on this database, potentially because respiratory rate in adults typically have smooth and slow waveforms, compared to the rapid peaks characteristic of heart sound signals (for which the MPT method was originally intended).

In conclusion, the results look promising for successful respiratory rate estimation from the 74 recordings analyzed, but it is still difficult to make generalized assessments on the data because of the small sample size. Additional work must be done on a larger

data set, so that the data can be divided into a training and test set. This is required so that the algorithms' performance on the data will be a better (unbiased) estimate of their actual performance in the field and avoid overfitting to the set of recordings in the test database. While there were other lung sound collections for auscultation training purposes, the databases were not pooled for the respiratory rate analysis because they would introduce many uncontrolled variables in the recording environment. The analysis would be adversely affected by differences between each database, including different recording equipment and different protocols on dealing with patients to capture their sounds. Instead, new data using the mobile phone recording environment (described in Chapters 4 and 5) will be collected for further algorithm development.

Chapter 4

Device Instrumentation

4.1 Hardware Issues with Recording on Mobile Phones

In recording heart and lung sounds on a mobile device to send to a specialist for interpretation, it is necessary to ensure that the recording is as faithful to the original sound as possible to preserve diagnostic information. In an attempt to reduce undesirable distortions in the audio signal, the phones with the most advanced sound card capabilities available on the market are used for this initial proof-of-concept work. The rapid evolution of the cellular phone industry should allow the technology available in high-end phones to become cheaper over time, for eventual inclusion in the rest of the phone models. In addition, the techniques developed for higher end phones can gradually be adapted to lower end phones with a systematic quantification of the degradation of clinical accuracy. Users can then choose to trade off cost for accuracy according to a given medical context.

Figure 4-1. Phones Considered for Research. From left to right:
Nokia N82, iPhone 3GS, HTC G1, and Motorola Droid.

The T-Mobile G1, Motorola Droid, Apple iPhone 3GS, and Nokia N82 phones were available for consideration for use in this research (see Figure 4-1). As an initial qualitative assessment, heart and lung sounds were recorded on the phones (see Figures 4-2 to 4-4) except for the Droid (because it was not released to the market until later in the project). In general, the recordings had poor quality because there was no stethoscope attachment used with the mobile phones to amplify heart and lung sounds, resulting in large noise spikes in the data. Some breaths were completely missed by the mobile phone, for example with the N82 phone in Figure 4-2. The tracheal sounds from the G1 in Figure 4-3 seemed promising, but the signal was soft (due to low amplitude), so there were still opportunities for improvement.

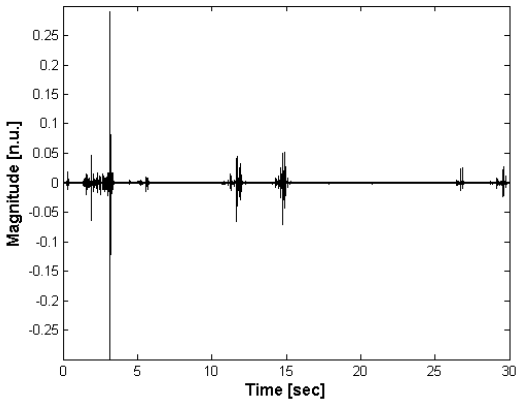


Figure 4-2. Nokia N82 tracheal breath sound recording.
From healthy female adult with respiratory rate 20 cpm.

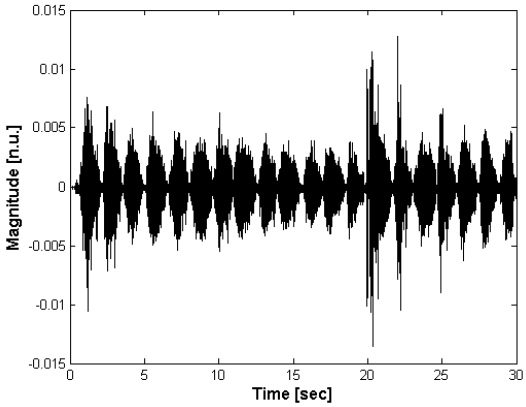


Figure 4-3. T-Mobile G1 tracheal breath sound recording.
From healthy female adult with respiratory rate 20 cpm.

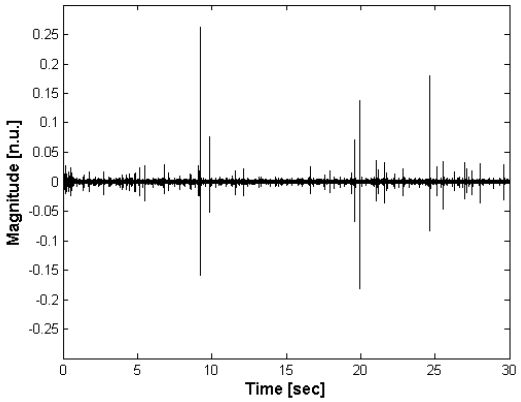


Figure 4-4. Apple iPhone heart sound recording at the base of the heart.
From healthy female adult with pulse rate 88 bpm.

From the preliminary tests, it appeared that there was good potential for capturing heart and lung sounds, but the system had to amplify the signals of interest to be more audible. One phone was needed to build this prototype system, and a more quantitative approach was taken to make this decision. The most important criterion was that the audio analog-to-digital converter (sound card) on the phone had to have a flat frequency response in the range of interest. From Chapter 1, the range of interest for heart sounds is between 2Hz to 150Hz, while lung sounds span from 100 Hz to 1200 Hz (Piinla 1995, Pasterkamp et al. 1997). A flat frequency response in a given frequency range indicates very little distortion for a signal in that range. A majority of phones, particularly cheap models, have poor frequency responses in these low frequency regions (GMSArena 2008a, 2008b, 2008c). In terms of audio response, the main motivation for designing the frequency response of mobile phone sound cards has been to provide intelligible data in the human speech frequency range; 200 Hz to 4 kHz (Cisco Systems 2007). However, since heart and lung sounds occur at much lower frequencies than voice, only certain phone models available on the market will perform well at capturing these vital signs. More recently, mobile phone manufactures have

improved the lower-end frequency response of phones to enable playback of high-quality music. The human ear is typically sensitive down to around 20 Hz, and so an increasing number of phones have been manufactured to provide undistorted frequencies as low as 20 Hz (or sometimes lower). A fortuitous consequence of this development is that sound cards, which need to be duplex to allow two-way conversations, allow high-quality recordings at diagnostic low frequencies.

GSMArena.com (2010) is an organization that gathers information on different manufacturers of mobile phones and posts detailed technical reviews on phones online. They perform thorough audio quality tests using the *M-Audio Fast Track Pro* hardware audio interface (M-Audio, Irwindale, CA) and *RightMark Audio Analyzer* software (Independent Audio Benchmarking Project 2009) to determine the frequency responses of the phones. More information on the test procedure can be found at GSMArena.com (2007a). The Nokia N82 exhibits a nonlinear frequency response characteristic below 500 Hz (GSMArena.com 2007b) where signals were increasingly attenuated from 500 Hz down to 20 Hz (the lowest frequency tested) (See Figure 4-5). As can be seen from Figure 4-6, the HTC G1 phone has a relatively good response for signals above 200 Hz but exhibits low frequency distortion below this frequency (GSMArena.com 2009). Figure 4-7 illustrates a comparison of the frequency responses of the Droid and iPhone 3GS (GSMArena.com 2010), where both have flat frequency response characteristics for input signals from 20 Hz to over 10 kHz, encompassing much of the frequency range of heart and lung sounds. Note that the frequency response of these phones below 20Hz is not available, so distortion characteristics below this frequency are unknown for each phone.

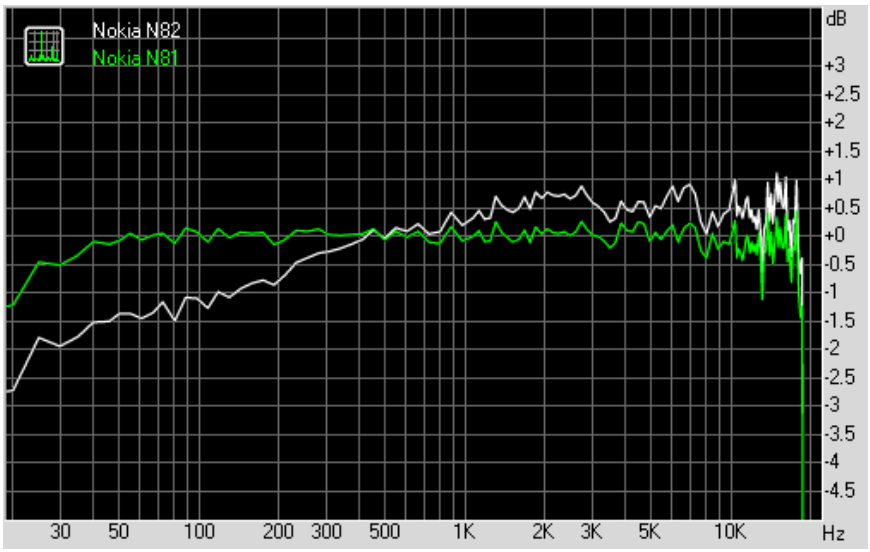


Figure 4-5. Comparison of the frequency responses of the Nokia N82 (white) and N81 (green) devices. Reproduced with permission from GSMArena.com (2007).

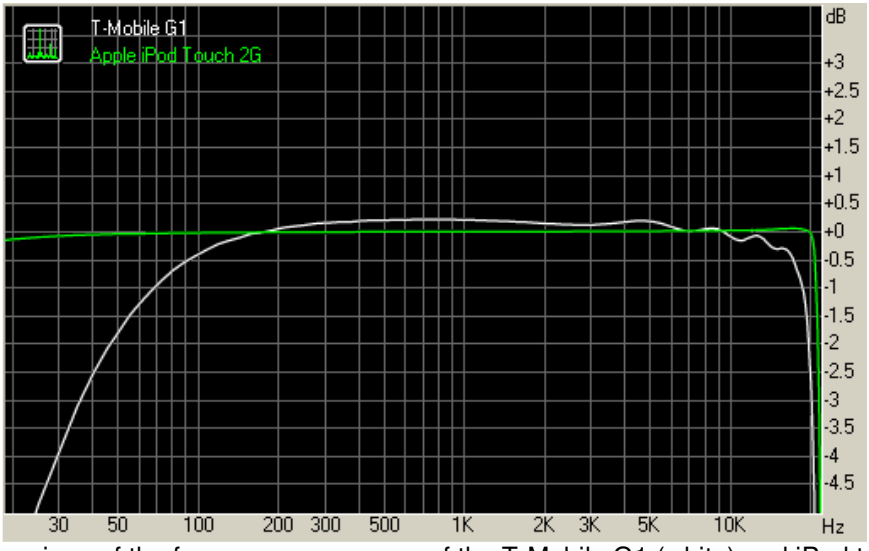


Figure 4-6. Comparison of the frequency responses of the T-Mobile G1 (white) and iPod touch 2G (green) devices. Reproduced with permission from GSMArena.com (2009).

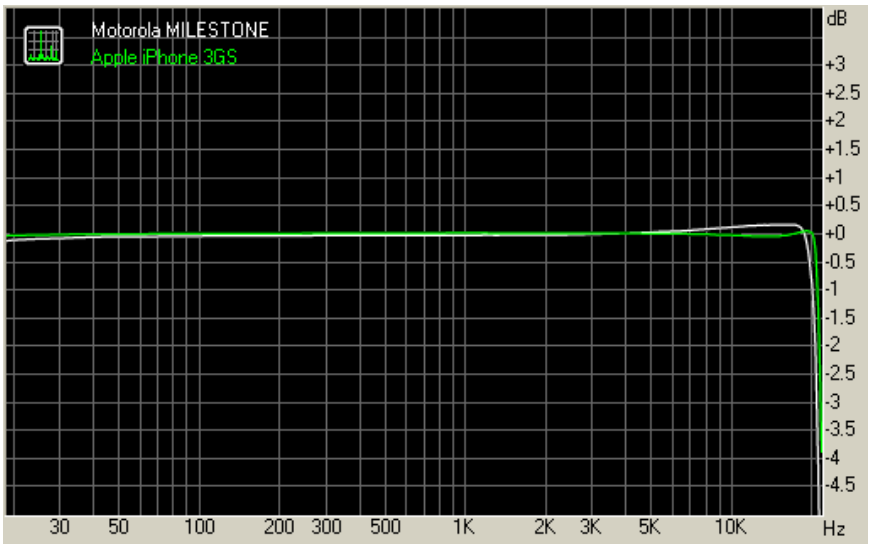


Figure 4-7. Comparison of the frequency responses of the Motorola Droid (white) and Apple iPhone 3GS (green) mobile devices. Reproduced with permission from GSMArena.com (2010).

Phone Model	Noise Level (dB)	Dynamic Range (dB)	Total Harmonic Distortion (n.u.)	Intermodulation Distortion & Noise (n.u.)	Stereo crosstalk (dB)
Motorola Droid (Milestone)	-74.1	74.7	0.014	0.040	-73.1
Apple iPhone 3GS	-92.1	92.1	0.0035	0.289	-86.1
T-Mobile G1	-87.2	89.2	0.017	0.289	-86.1
Nokia N82	-81.6	80.8	0.0046	0.137	-69.3
Nokia 3120 Classic	-76.6	79.0	0.018	0.439	-73.9
LG KU990 Viewty	-86.8	83.8	0.020	0.124	-86.6
HTC Touch Pro	-85.3	87.9	0.027	0.267	-86.3

Table 4-1. Comparison of the audio quality tests on the candidate phones. Reproduced with permission from GSMArena.com (2007a, 2010).

Table 4-1 provides further details on the audio quality of each of the phones and some comparable alternatives. The noise levels provide information on the ratio between background noise and the signal of interest (i.e. the voice of an individual during a call) in the audio reproduced by the mobile device (GSMArena.com 2007a). Meanwhile, the dynamic range describes the difference in decibels between the loudest and softest sounds that can be captured by the phone at the same time. Then there is dynamic range, which is important in noisy environments because the body's sounds are typically much quieter than ambient background noise and still need to be heard in

the recordings. From Table 4-1, the Apple iPhone 3GS has the largest dynamic range at 92.1 dB. The total harmonic and intermodulation distortion values convey the amount of distortion that occus between the frequency information in the original input signal and the one reproduced by the phone (Carter and Mancini 2009, Anand 2006). The last column in Table 4-1 displays stereo crosstalk, which is the amount of sound leaked into another channel when an input signal is applied to a different stereo channel (Graf 1999). These audio quality measures are all important quantitative ways of evaluating the ability of these mobile devices to record heart and lung sounds accurately.

The phones were compared based on these audio quality test results and device features in order to select the optimal mobile phone for use in testing and data capture. The Nokia N82 phone had moderate noise level, dynamic range, and distortion compared to the other phones, but a poor frequency response curve. Another disadvantage was that the Nokia app store (for finding recording software) was only announced in May 2009 (several months before the start of this thesis work) and was not as mature as the app stores of Apple or Android (Woyke 2009). At the time, the only way to record audio on the N82 was through the multimedia messaging capability, which would save the recording as an AMR file using an unknown compression algorithm. Alternatively, the G1 phone had a better frequency response curve than the N82 and had access to the Android Market for applications. However, the frequency response curve for the G1 still low pass filtered crucial signal content. The phone also used the Android 1.6 OS (and had not been updated over the air yet), while other Android phones were already using the more robust OS versions 2.0 and 2.1. Furthermore, the G1 phone's 11-pin mini USB port also served as its audio jack (HTC

2009), so the number of compatible hands-free kits were limited to those with the same port (or required the purchase of an extra adapter). As mentioned above, the Motorola Droid and Apple iPhone were found to be most ideal for capturing heart and lung sounds because of their flat frequency response curves. They both also had standard 3.5mm audio jacks, which allowed use with a wider range of headsets than the G1 phone. The iPhone had slightly better performance than the Droid on audio quality tests presented in Table 4-1. However, the Droid ran the Android operating system, which meant that it could be integrated with the Sana platform (currently only supported on Android devices). This was a large advantage for the Droid over the iPhone because it would allow capture of audio for transmission to a medical record system via intermittent connectivity over WiFi/GPRS/3G. Being able to record heart and lung sounds on the iPhone without infrastructure to transmit the audio to a remote location would not offer much benefit to a rural healthcare worker, who would not need to record the sound if it was only used locally. As a result, the Motorola Droid was selected for development and testing in this thesis.

Even though the hardware is capable of high quality recordings, caution must be used when capturing audio data. The audio channel on the device (for regular mobile phone voice calls) cannot be used for transmitting heart and lung sounds because telecommunication companies often selectively filter call signals to eliminate noise from the voice conversation. Signals from 300 Hz to 3 kHz can typically pass through the transmission networks; but even those signals are often compressed, which would likely distort or filter out heart and lung sounds (Cisco Systems 2007, Istepanian et al. 2006). To provide some perspective, it is estimated that energy in 95% of S1 and 99% of S2

heart sounds occur at frequencies less than 75 Hz, so cellular communication networks would not faithfully transmit these heart sounds (Istepanian et al. 2006). There have been some attempts at complex compression methods using wavelet transforms or fractal dimension analysis, but none have been accepted as a standard in compression for heart and lung sounds nor been adopted widely (Hadjileontiadis et al. 2002, Hadjileontiadis and Rekanos 2003). Without knowledge of a robust compression algorithm that does not distort clinically useful information, it is necessary to store the data on the phone in an uncompressed and lossless method in a standard binary format (such as the WAV file format). The WAV file format is limited to a size of 2 GB because of the 32 bit header that stores the file size (IBM and Microsoft 1991, Microsoft 1994). However, longer duration recordings can easily be saved by removing content at frequencies higher than 4 kHz, which is well beyond the range of clinical interest. A recording at 8kHz/16 bit with uncompressed (pulse code modulation) WAV encoding equates to 128 kbit/s and a file size of 54.9 MB/hour. This would allow a file to be as long as 37 hours. Traditional auscultation by a doctor on a patient usually involves listening for only several breaths at various locations along the back and for several heartbeats along the heart, so several hours of recordings time would be more than sufficient.

The Rehearsal Assistant application (UrbanSTEW 2009), which can be downloaded for free from the Android Market, was used for recording the heart and lung sounds on the Droid phone. Within the application's settings menu, the "Record Uncompressed" option was enabled to record the audio in an uncompressed WAV

format on to the phone's micro SD card. The files could then be transferred easily to a computer for offline analysis.

A hands-free kit (comprising a low-cost earpiece, microphone, and connector to the phone) was used to allow for greater flexibility in positioning the microphone on the body for capturing the loudest signal possible. Without a hands-free kit, placement of the mobile phone directly on the body performed poorly because of the lack of an airtight seal between the body and the microphone. Three hands-free kits (displayed in Figure 4-3) were tested. The Droid phone has a standard 3.5mm audio/microphone jack, so there were actually many other available options from Motorola, third-party vendors, and even Apple.

(a) (b) (c)

Figure 4-8. Different hands-free kits tested with the Droid phone.
The devices are from a) Motorola, b) Verizon Wireless, and c) Apple.

After numerous trials on a healthy female test subject, the use of a hands-free kit (directly on the body or attached to a cup) was not able to audibly record heart or lung sounds. Supplementary materials were necessary to acoustically amplify the signal from the patient. As mentioned in Chapter 1, midwives in rural locations will often hold a cup up to their ear and press it against the abdomen of a pregnant woman to listen to the fetal heart beat. A natural design for a low-cost stethoscope is then to puncture the base of the cup, attach the microphone of the hands-free kit, and record the acoustic data captured by the cup. Test recordings were attempted with cups made of different

materials, such as metal, paper, Styrofoam, and plastic (see Figure 4-9), to examine the sound propagation qualities of each material.

Figure 4-9. Cups of different material for testing with the hands-free kit.

The decision on which hands-free kit to use did not come until after the entire prototype (described later in Section 4.5) was finished because the sounds were not audible enough for comparison otherwise. Using the final stethoscope prototype design with the Droid phone, the three hands-free kits were tested. Although there was largely no noticeable difference to the human ear upon playback of the recordings, the signal waveforms of the heart and lung sounds in Figures 4-10 and 4-11 were more helpful in distinguishing the performance of the kits from each other. The Motorola headset (top row in Figures 4-10 and 4-11) seemed to exhibit the poorest quality because of its susceptibility to noise as evidenced by the large noise spike in the top left plot of Figure 4-10. The lung sound waveform for this headset also demonstrated unwanted noise disturbances for almost every breath cycle. This performance suggested the importance of the shape of the microphone. The microphone of the Motorola headset (left-most kit in Figure 4-8) was mounted on a curved surface of the device, which was difficult to attach securely to other materials or flat surfaces, creating sound leakage from the body. On the other hand, the Verizon headset (middle row in Figures 4-10 and 4-11)

provided a less noisy heart sound signal, but the 10-second interval of the heart sound showed multiple instead of singular spikes for each S sound. The kit that delivered the cleanest signal was the iPhone Stereo hands-free kit (bottom row of Figures 4-10 and 4-11). There were not any large noise peaks or asymmetries in the plots; the signals followed consistent sinusoidal patterns while the heartbeats and respiratory cycles could be easily distinguished. (More discussion on the signal characteristics of the recordings from the Droid, iPhone headset, and stethoscope device combination are found in Section 4.5). This hands-free kit also had a flat microphone piece that could easily be taped to different surfaces, such as a cup. The tight coupling of the microphone to any other supplementary material would prevent interference from ambient noise.

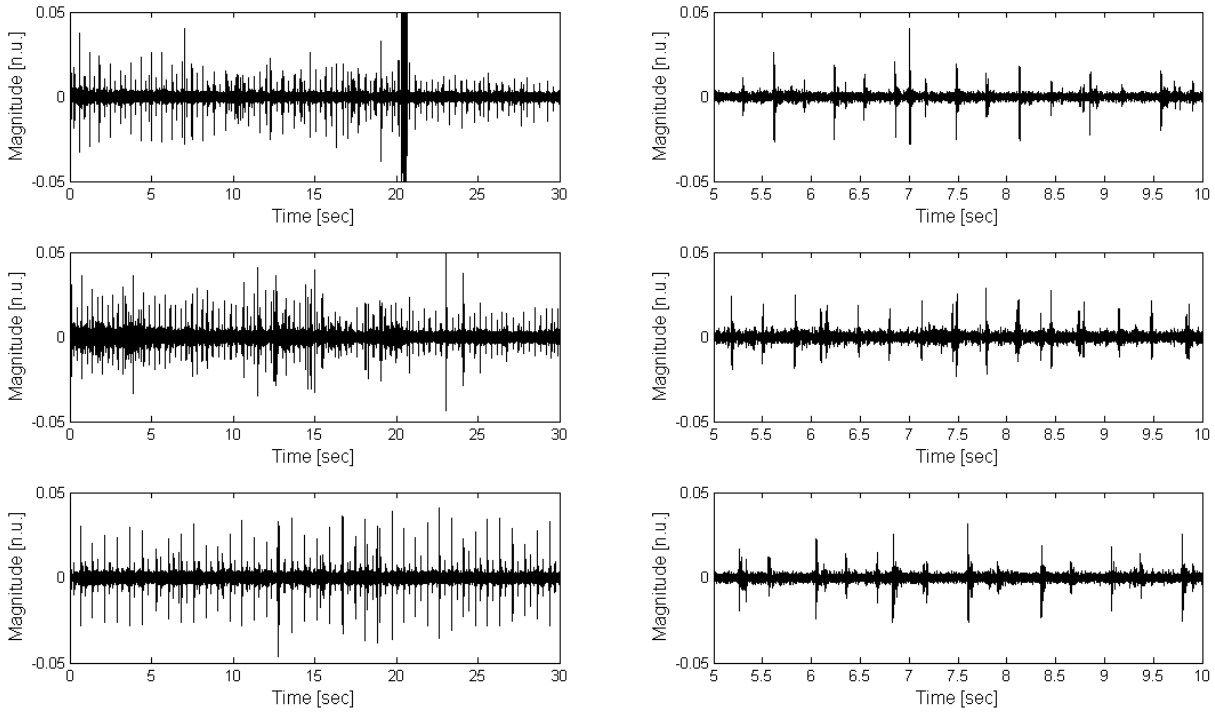


Figure 4-10. Heart sound recordings with Droid phone and stethoscope device with different hands-free kits: Motorola (top row), Verizon (middle row), and Apple (bottom row). The left column is the full 30-second recording, while the right column shows only a 10-second segment of the full recording.

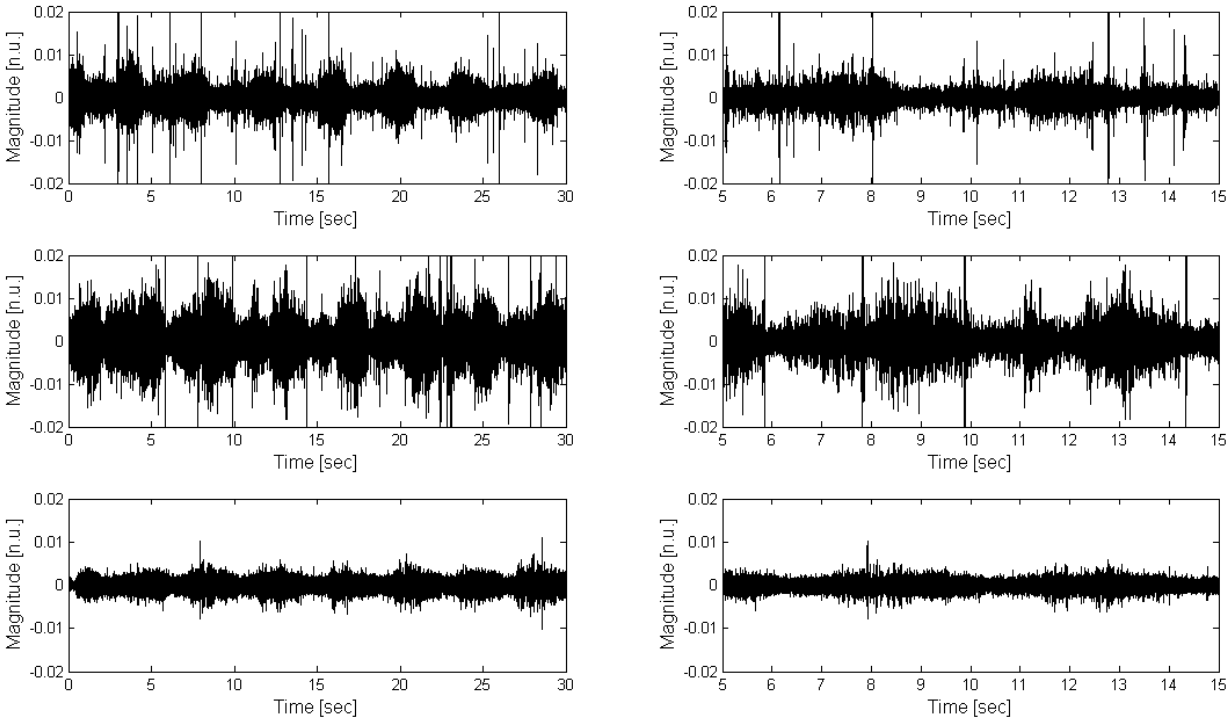


Figure 4-11. Lung sound recordings with Droid phone and stethoscope device with different hands-free kits: Motorola (top row), Verizon (middle row), and Apple (bottom row). The left column is the full 30-second recording, while the right column shows only a 10-second segment of the full recording.

The reasons for selecting the Droid phone and iPhone hands-free kit have been explored, and the following sections describe the prototype process of developing the stethoscope attachment. To achieve the desired signal amplification on with the mobile phone, commercial stethoscopes were examined to discover how they are able to successfully amplify body sounds. The historical development and acoustic properties of stethoscopes are presented in Section 4.2.

4.2 Stethoscope Background

Stethoscopes are used for performing auscultation, which is the act of listening to the body's internal sounds. When a patient breathes or their heart beats, there are small vibrations throughout the body that are picked up and amplified by the stethoscope's

diaphragm, which travel through hollow air-filled tubes to two earpieces for a doctor to listen through (Scanlon and Sanders 2005). Doctors can use a stethoscope to listen to a patient's lung sounds and detect abnormal conditions including asthma, pneumonia, and bronchitis. Stethoscope can also be used for listening to heart sounds for detection of murmurs or diseases such as heart failure.

The stethoscope was first developed in 1819 by the French doctor Rene Laennec when examining a patient (Sheldon and Doe 1935). He rolled up many pieces of paper into a sturdy roll with one end placed on the patient's chest and one end up to his ear - making the first monaural (single earpiece) stethoscope. Laennec later refined the design into a foot-long wooden stethoscope with a funnel-shaped cone at one end. Then George Cammann invented the binaural stethoscope in 1843, followed by Dr. Littman's stethoscope in the 1960s (3M).

Today, acoustic stethoscopes are comprised of the following parts: earpieces (made of rubber to reduce outside noise), ear tubes (made of hollow metal tubes), acoustic tubes (made of rubber to connect ear tubes to the chest piece), and a metal chestpiece (one or two-sided with a bell and/or diaphragm) (Trimline Medical Products 2004). An example of an acoustic cardiology stethoscope is shown in Figure 4-12.

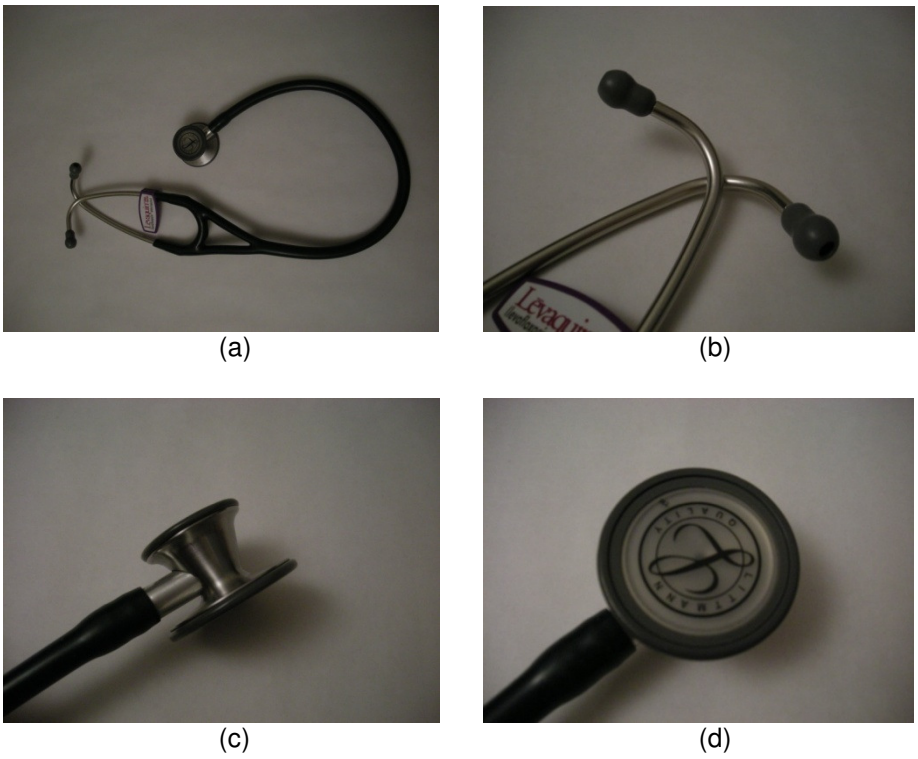


Figure 4-12. Littmann Cardiology III Stethoscope which cost $150 or more. a) entire device, b) ear pieces, c) side angle of the chestpiece with diaphragm and bell sides, and d) chestpiece diaphragm side.

Doctors also use electronic stethoscopes, which operate by converting the body's acoustic signals into electrical ones (WinHealth 2010). Electronic stethoscopes are useful for amplifying heart/lung sounds, while minimizing background noise using in-built signal processing or hardware filters. They typically aim to have flat frequency response curves for signals in the range of 20 Hz to 5 kHz (Istepanian et al. 2006). Some models also have additional functionality such as patient heart rate estimation, visual displays, record and playback capabilities, and the ability to transmit data wirelessly. When using electronic stethoscopes, a concern is that diagnostically relevant components of the signal may be distorted or attenuated.

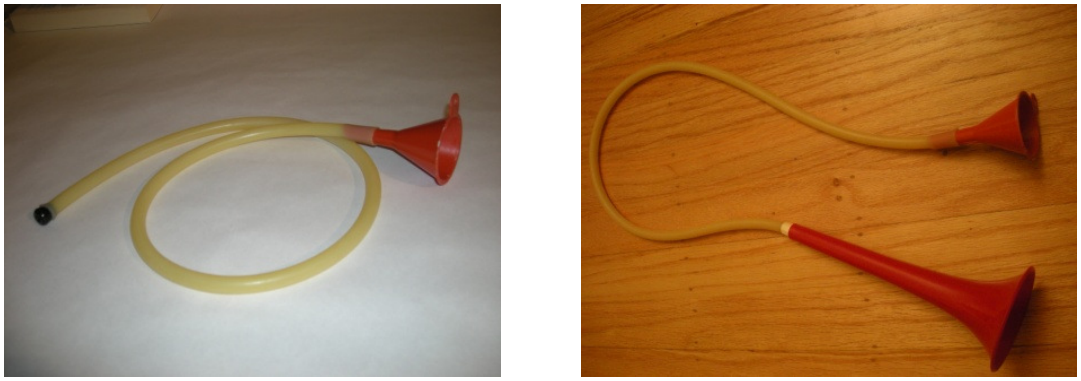
Both acoustic and electronic stethoscope devices can be quite expensive. High quality stethoscopes cost at least $150 and cannot be distributed to thousands of rural

healthcare workers in a cost-effective manner (WinHealth 2010). In addition, auscultation requires specialized training which is hard to provide for basic healthcare workers. Alternatively, there are cheap, one-time-use, disposable stethoscopes that U.S. hospitals bulk-purchase for use in the intensive care unit (ICU). Patients in the ICU are especially prone to infection, and the same stethoscope on multiple patients should not be used since the placement of them on the patient makes them an obvious disease vector. However, from anecdotal evidence, doctors do not recommend these cheap stethoscopes because the devices do not amplify the patient's biological sounds sufficiently for accurate diagnosis. Using already-manufactured stethoscopes to capture heart and lung sounds with a mobile phone is also not a sustainable solution because of the reliance on a steady supply chain of devices and replacement parts. Instead, a simplified version of a high-end stethoscope that is prototyped with locally available parts is preferable. The prototype design process occurred in two parts. The first phase involved understanding the important functions of different parts of the stethoscope and the building of a prototype able to record heart sounds (see Section 4.3). The second phase involved using a larger variety of materials to discover which combinations would amplify sounds the loudest and allow capture of both heart and lung sounds (See Section 4.4).

4.3 Design Phase 1: Stethoscope Device Prototype for Mobile Phones

Educational material about stethoscopes contains descriptions on how to construct a homemade stethoscope, such as from two plastic funnels, one foot of old

rubber garden hose, and some glue or tape (Science Fair Adventure 2007). A vinyl tube or paper towel roll could also be used instead of the rubber hose, and the funnel could be replaced by a water bottle that was cut in half (George 2009). The online material suggested different adhesives including clay, caulk glue, electrical tape, and duct tape. By following these instructions, several initial prototypes were created in Design Phase 1 (See Figure 4-13).



(a) (b)

Figure 4-13. Initial stethoscope prototypes made from rubber tubing and plastic funnels.

Using a funnel is a straightforward way to acoustically amplify sound waves. The mouth of the funnel should be applied against the patient's body with enough pressure to create an airtight seal between the body and the funnel. The beating of a person's heart, or the act of breathing, creates physical vibrations that travel to the surface of the body. When a funnel is placed over a patient's heart or lungs, the body's physical vibrations translate into air vibrations, which enter through the top of the funnel. Because of the large difference in diameter between the top and base of the funnel, small vibrations that enter through the top of the funnel are magnified into much larger vibrations when the air attempts to exit through the base of the funnel. As the air travels through the tube and up to the ear, the air vibrations that hit the eardrum are interpreted as audible sound by humans. The larger the air vibrations, the louder the perceived

sound, so the best stethoscope devices must be able to induce large air vibrations. One factor that contributes to the performance of the device is the ratio between the two diameters of the funnel, which directly determines its degree of sound amplification (see Figure 4-14).

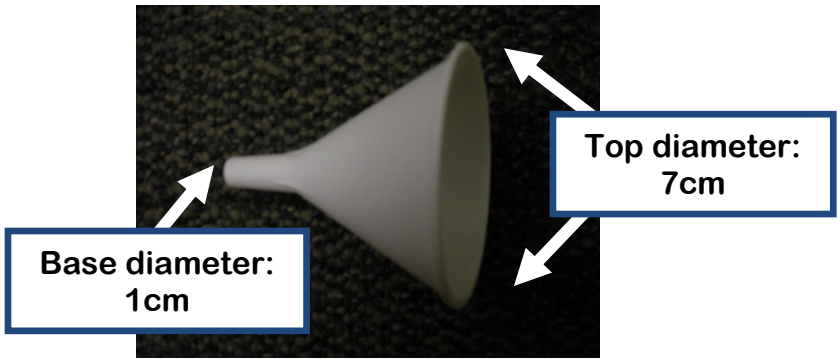


Figure 4-14. Example of a plastic funnel.

Another method to amplify sound is to add a diaphragm to the funnel, where the diaphragm is a piece of material that is held taut such that it vibrates when it is hit. The diaphragm concept is similar to the membrane (or skin) fitted over the musical drum instrument, which resonates when the surface is struck. When a diaphragm is stretched over the mouth of the funnel, the body's physical vibrations directly come in contact with the membrane, which will push the air behind the membrane back and forth more forcefully (than compared to no diaphragm present). The larger vibrations then induce a louder signal. In initial experiments, a latex glove was used as a diaphragm over the mouth of the funnel and secured with masking tape (see Figure 4-15a). The glove was successful in creating a taut diaphragm for the device and resonated audibly when tapped very lightly. In another alternative, a piece of tougher plastic, which was cut from leftover plastic packaging, was secured to the top of the funnel with scotch tape, but did not vibrate very well during testing (see Figure 4-15b).

(a) (b)

Figure 4-15. Addition of diaphragms to experimental prototypes.

A great deal of noise from the environment was inadvertently recorded during tests with the prototypes. As a solution, a small hand towel was wrapped around the funnel in order to block out ambient noise. Because the towel was made of soft material, it would absorb outside noises, instead of letting the noise reach and interfere with the sound captured by the funnel. At the end of the Design Phase 1, the resulting prototype (as shown in Figure 4-16) was constructed from a plastic funnel, latex glove, vinyl tube, earpiece, thin black rubber material, masking tape, and a hand towel.

(a) (b)

Figure 4-16. Prototype at the end of Design Phase I

Three licensed medical doctors reviewed a collection of heart and lung sound sounds recorded on an Android G1 phone with the Design Phase 1 prototype. The heart

sound recordings were determined to be of sufficient quality for detecting murmurs, arrhythmias, and a third heart sound. In the abnormal cases though, the subjects would need to be referred to a heart specialist for further evaluation. Unfortunately, the lung sounds were too soft to be able to make a confident diagnosis. Another drawback of the prototype was the inconvenience of pressing the end of the vinyl tube against the microphone in a firm manner, while simultaneously holding the funnel up against the patient's body. Because of these limitations, there was a need to iterate on the design. The new design needed to improve on how the acoustic device would attach to the hands-free kit without depending on too much effort from the user. The goal was to develop a *single* device for recording both heart and lung sounds. Requiring a healthcare worker to use two devices (one for heart sounds and one for lung sounds) would add unnecessary complexity to the workflow.

4.4 Design Phase 2: Iteration on Stethoscope Prototype

The second design phase involved the use of a larger variety of materials and collaboration with individuals in areas of expertise such as low-cost medical diagnostics, doctors, and engineers. By building off of the Design Phase 1 prototype, new materials were substituted for individual parts of the prototype to test for incremental improvements in the performance of the device. The new materials were still low-cost household items that could be purchased from local hardware or convenience stores as displayed in Figure 4-17.

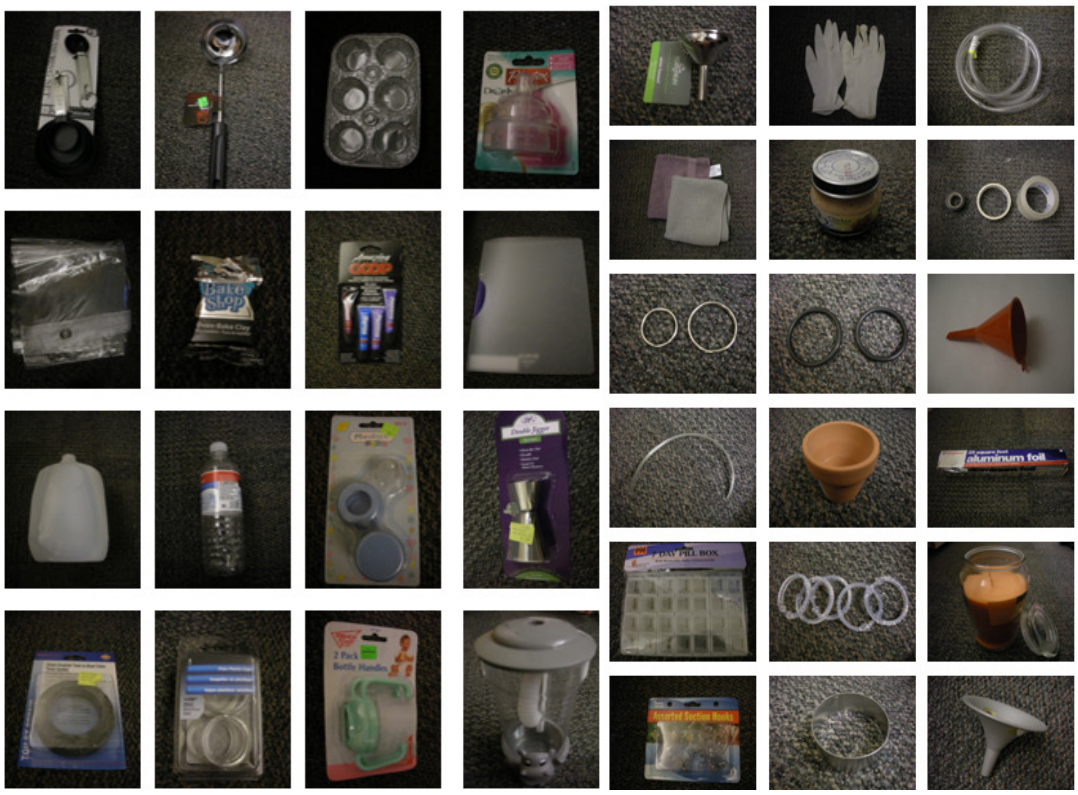


Figure 4-17. Household items used for prototyping a stethoscope attachment for a mobile phone.

Several of the materials purchased needed some minor modifications, which were performed at the MIT Edgerton Center Student Shop and Central Machine Shop. The primary use of the machine shops were to drill holes in many of the components, so the sound would be captured by the component and enter the microphone of a hands-free kit through the hole (see Figure 4-18).

Figure 4-18. Household items with holes drilled through them for the hands-free kit microphone.

A compilation of all the resulting prototype designs, made from the household items, are shown in Figure 4-19. The “chest piece” of the stethoscope was replaced with a plastic funnel, water bottle, baby food jar lid, candle holder, muffin tin, measuring spoon, candle jar lid, shot glass, or soup ladle in the new prototypes. The “diaphragm” was replaced with a circular cut-out from an overhead transparency, plastic folder, shower curtain, or piece of plastic packaging. The diaphragm was attached securely to the chest piece in a number of methods using an O-ring, super glue, rubber gasket, or clay. The chest piece typically had a hole on the other end such that the sound could enter the microphone of the hands-free kit.

Figure 4-19. Stethoscope prototypes made from household items.

One design challenge was figuring out how to affix the microphone of the hands free kit to the physical stethoscope prototype. Various methods for holding the device included using Styrofoam, Velcro, tissue paper, aluminum foil, and masking tape (see Figure 4-20). Masking tape proved to be the most effective and convenient method because it was flexible enough to accommodate any microphone shape, as long as the seal made by the tape was airtight. The aluminum foil added extraneous noise to the recordings when crinkled, while the other methods (Velcro, Styrofoam, small box) did not provide an airtight seal.

Figure 4-20. Methods for securing microphone to the prototype device.

As the prototypes continually evolved throughout the design process, we sought input from others who had more experience in the realm of low-cost diagnostics for resource-poor communities. In collaboration with the instructors of the MIT D-Lab Health class, José Gómez-Márquez and Amit Srivastava, a lab activity was developed for the students. The background context and motivation for creating a low-cost stethoscope to record heart and lung sounds was explained to students. Then they were split into groups to find a solution to the design problem. Each group was provided

with $20 worth of materials similar to those pictured in Figure 4-19, and they had a week to come up with a viable prototype. The assignment also required them to submit a lab report describing the prototype and design decisions, sample audio files recorded with their prototype, and images from simple signal processing on the recordings.

The students presented their creative designs to the class and a panel of guest judges including a pediatric pulmonologist, mechanical engineer, and medical student. The results of their work are portrayed in Figure 4-21. After testing many different materials (ceramic, steel, and plastic) for a “chest piece”, a majority of the groups found that stainless steel worked the best in amplifying sound. For the “diaphragm”, one design involved the use of a leather glove stretched over the mouth of a funnel, while another group used a latex glove with a piece of plastic taped on top for increased stiffness. Several groups used a vinyl tube to attach the funnel to the microphone of a hands-free kit. One group even inserted a tube within another tube for reduced interference from outside ambient noise. In order to attach the hands-free kit to the stethoscope they developed, one group capped the end of the vinyl tube with a rubber stopper. (There was a perforated hole in the center of the stopper so that the sound could enter the hands free kit.) Another group drilled a hole in a plastic pill box to insert over the base of a funnel. Then the microphone of the hands-free kit rested inside the closed pillbox, which was padded with extra tissue paper.

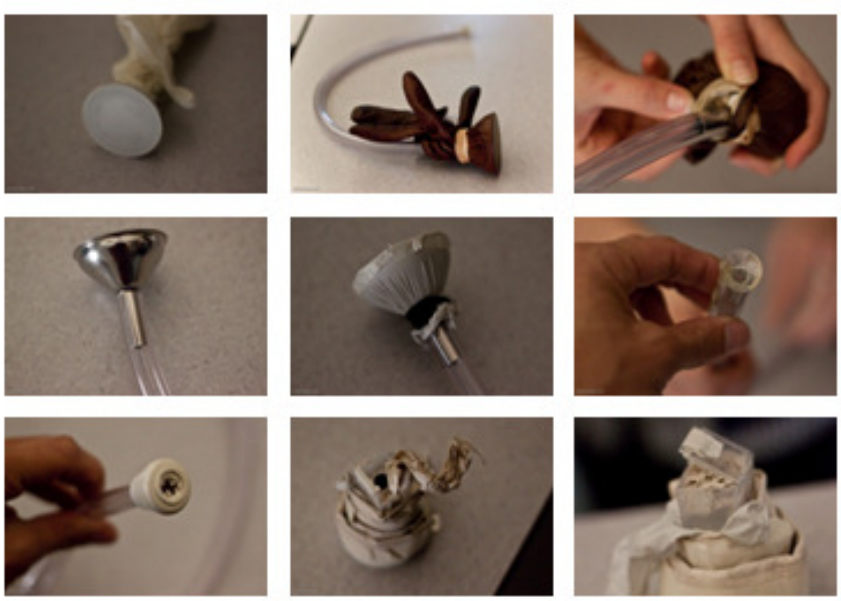


Figure 4-21. Stethoscope attachments for a mobile phone, designed by D-Lab Health students. Photos taken by Nathan Cooke.

To demonstrate the effectiveness of their designs, the students recorded heart and lung sounds and viewed the signal waveforms in MATLAB. Figure 4-22 shows the signal waveforms of a heart and lung sound recording (each with sampling frequency 44.1 kHz) from a stethoscope design consisting of a metal funnel, plastic overhead transparency, plastic tubing, and rubber doorstopper. The D-Lab students who created this prototype were Kristin Kuhn, Michael Melgar, Sudha Guttikonda, Kathleen Li, and Sivikami Sambasivam. Qualitatively speaking, the heart and lung sounds are audible in the recordings, and the heart sound signal waveform in Figure 4-22a shows distinct spikes for each heartbeat. However, the lung sounds waveform in Figure 4-22b does not show distinct respiratory cycles. This may be caused by the constant level of background noise heard in the audio recording, which ended up masking the respiratory phases.

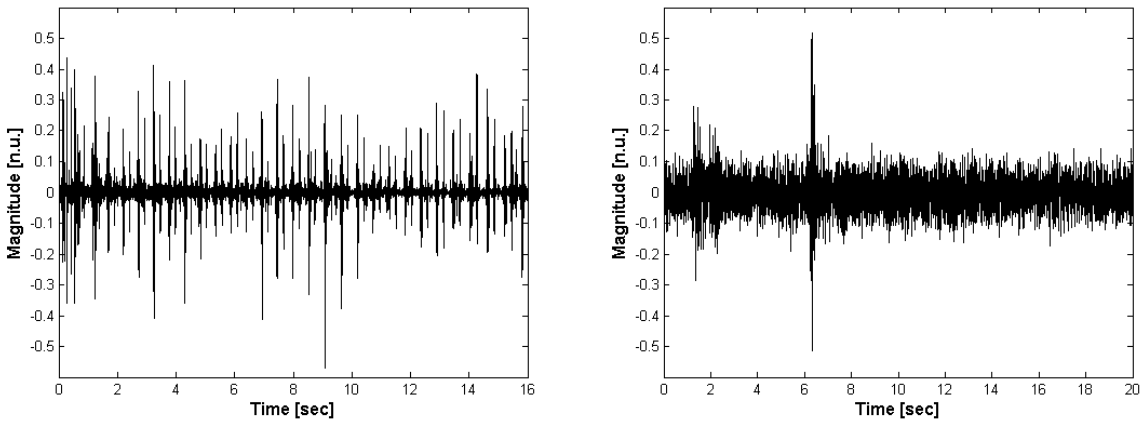


Figure 4-22. Sample heart (a) and lung (b) sound recordings
from a stethoscope attachment for a mobile phone made by a D-Lab student group.

From further research about how stethoscopes operate, in combination with the results learned from the D-Lab students' work, the following suggestions can be made for solutions to this design problem. The chest piece should be made of stainless steel or titanium for high-quality devices because sound is not as easily attenuated by dense metals (Stethoscope Basics 2010). Conversely, materials such as aluminum and plastic are used for making cheaper stethoscopes, but their performance is limited because sound is easily absorbed by soft materials (including light metals, plastics, and flexible rubber). The chest piece is typically 40-50 mm in diameter on most modern stethoscopes (Kazama 1990). There should be an airtight seal between the body and the chest piece in order for the sound vibrations to travel directly into the microphone of the hands-free kit, instead of being dissipated in the air outside the device. Additionally, a diaphragm made of thick plastic (i.e. polyethylene[32] compound film) that does not bend easily is suggested to induce stronger vibrations between the body and the microphone (Stethoscope Basics 2010). Tubing, which is often made from

[32] Polyethylene film is cheap, waterproof, flexible, thick, and can be purchased at local hardware stores (Popular Science 1959).

polyvinylchloride (PVC), can be added so that the sounds from the patient's chest can travel to a mobile phone resting on a table (instead of requiring the CHW to hold up the mobile phone in the air while being steady). Unfortunately, tubing can cause loss of signal quality if the sound must travel a long distance. In general, thick tubular walls are preferred because of a reduced risk in picking up ambient noise during sound transmission (Kazama 1990). These conclusions were incorporated into the process of designing the final prototype.

4.5 Final Design

The final design was approved by two professional medical doctors individually after listening to heart and lung sound recordings captured with the stethoscope prototype and mobile phone. The prototype is seen in Figure 4-23. The design includes a metal soup ladle (with the handle removed and a hole drilled in the center of it), a rubber gasket for toilet plumbing (with a slit cut around its circumference), and a piece of thick plastic cut in the shape of a circle. The plastic is used as a diaphragm by inserting it into the slit of the rubber gasket (made with a knife). Then the gasket is pushed into the soup ladle for a snug fit. Using masking tape, the microphone of the iPhone hands-free kit is attached to the soup ladle in the location where the hole was drilled. A hand towel is folded and used to cover the microphone from ambient noise during recording. The hands-free kit is plugged into the Motorola Droid phone, and sounds are captured with the use of the Rehearsal Assistant Android application with the "Record Uncompressed Audio" setting enabled (UrbanSTEW 2009).

The components in this stethoscope prototype are intended to be sufficiently modular such that the parts can be substituted for materials that are locally available in the deployment sites, making construction and maintenance simple and rapid. For example, the soup ladle can be replaced by a funnel. The diaphragm can be cut out of old plastic packaging, and the gasket can be substituted by a round piece of rubber or O-ring. As a rough estimate, the cost of this physical device is $4 ($1 for the rubber gasket, $1 soup ladle, $1 plastic folder, $1 hand towel) plus the cost of the hands-free kit, which can range from $20-$35 per device, but can be substantially less, especially if purchased secondhand.

Another advantage of having the device easily disassembled (see Figure 4-23a and c) is that it can be cleaned to prevent spread of disease from one patient to another. Wiping the plastic diaphragm and rubber rim of the device with rubbing alcohol can be done in between patients, as doctors normally practice for cleaning their traditional stethoscope chest pieces.

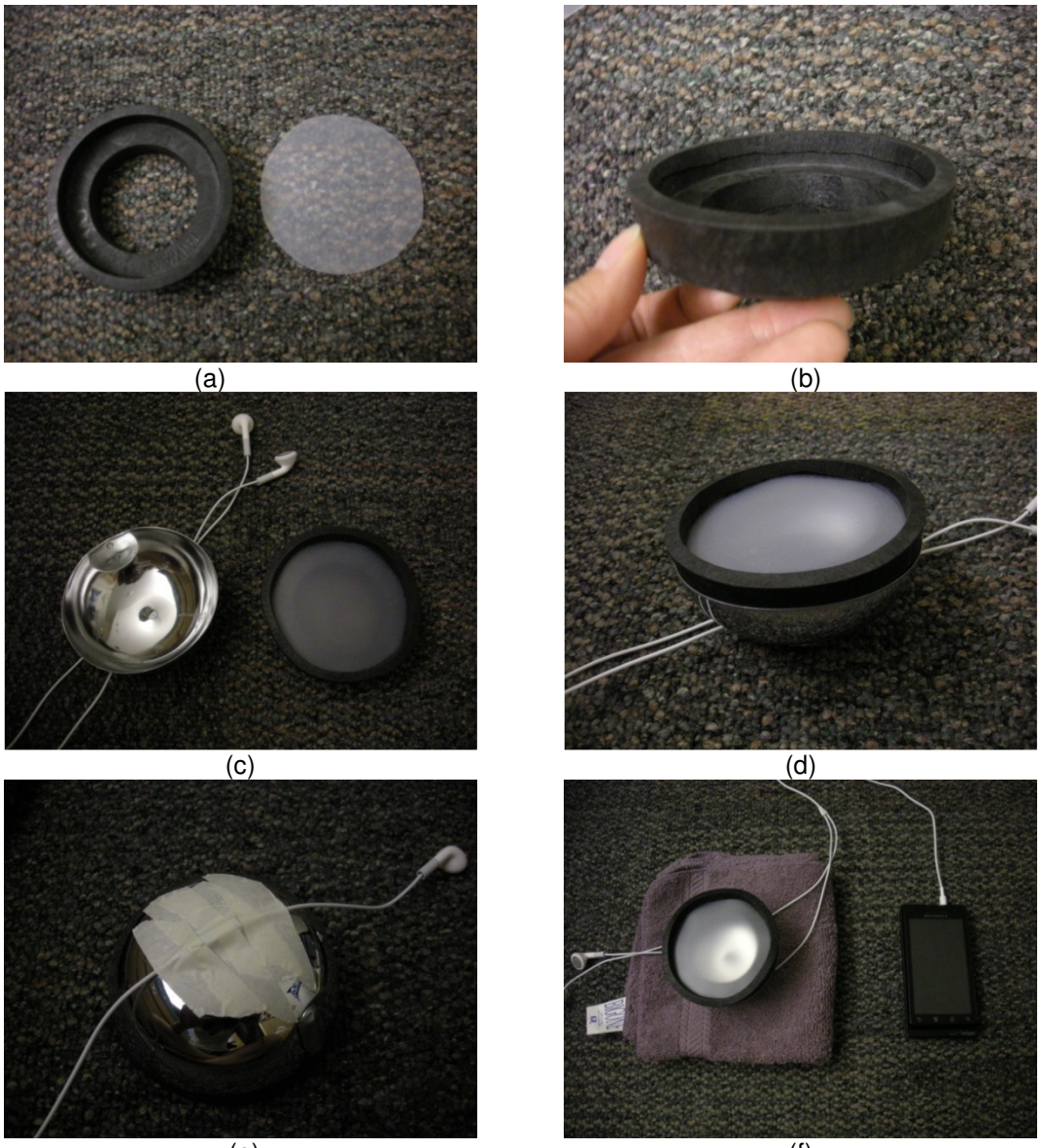
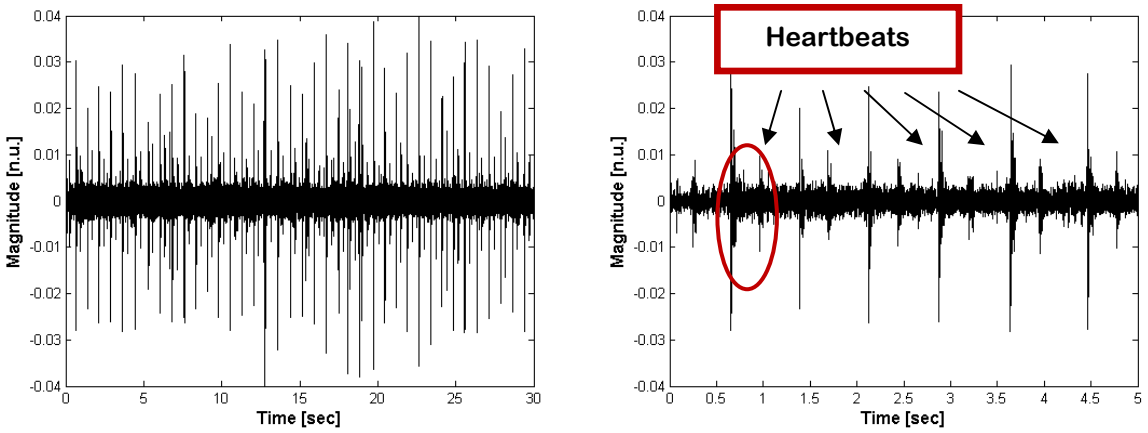


Figure 4-23. Final stethoscope attachment design for a mobile phone.

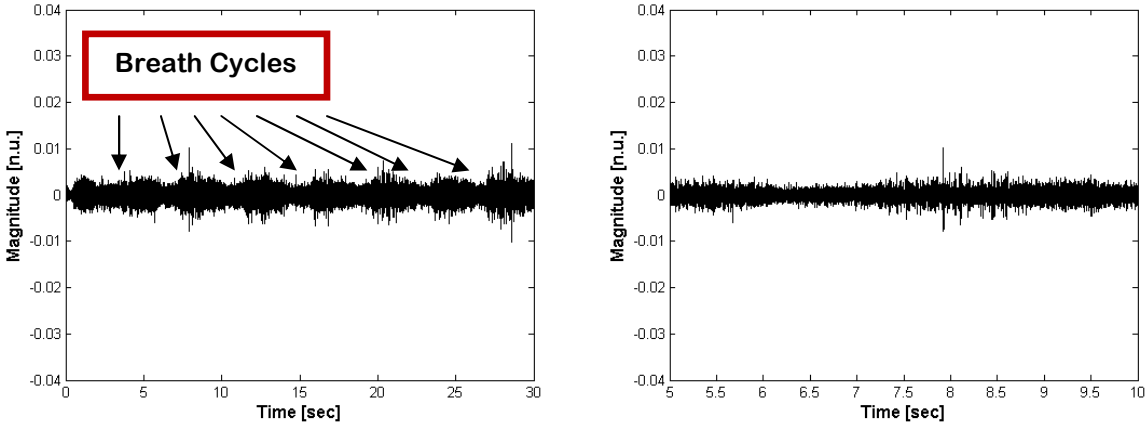
The stethoscope prototype is able to capture heart and lung sounds successfully, and the waveforms of two sample recordings (with sampling frequency 44.1 kHz) from a healthy female adult are displayed in Figures 4-24 and 4-25. Part ‘a’ of both figures show full waveforms of the 30-second recordings, while part ‘b’ offers a closer look at the waveforms over a five-second interval. The high frequency spikes of the heart beat are clearly seen in Figure 4-24a and b (6.5 heart beats can be counted in Figure 4-24b

as part of an overall heart rate of 88 bpm). There is also some slight respiratory amplitude modulation of the heart sound signal in Figure 4-24a, presumably from the chest moving towards and away from the microphone during breathing. The breath cycles are also evident in the lung sound signal of Figure 4-25a (7.5 breath cycles can be counted for a respiratory rate of 15 cpm). It can also be observed from comparison of the signal amplitude in Figures 4-24 and 4-25 that sounds at the apex of the heart are more audible than vesicular lung sounds, as claimed in Chapter 1.3. Further discussion on the use of algorithms to extract respiratory and heart rate from the patient recordings is offered in Chapter 5.



(a) (b)

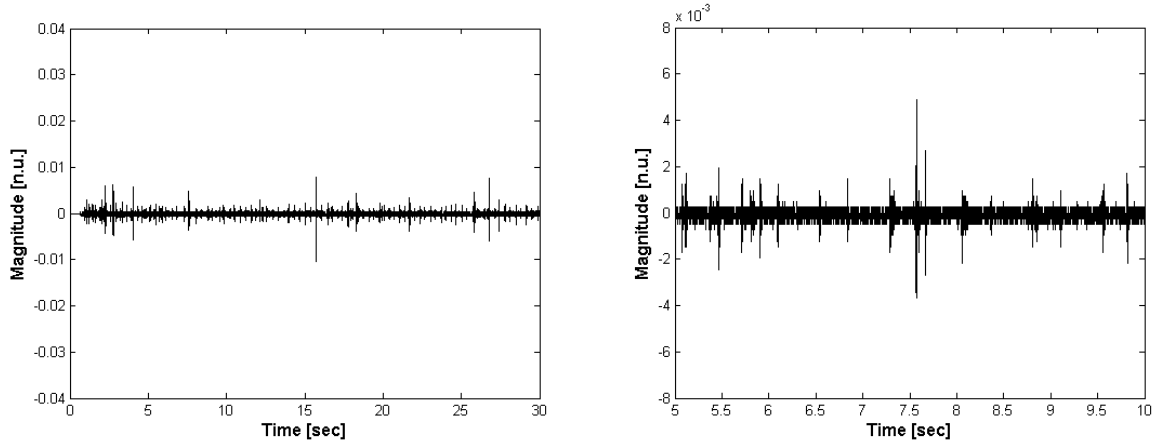
Figure 4-24. Uncompressed heart sound recording from healthy female adult using stethoscope device attached to Droid phone. Full 30-second signal (left) and five-second segment (right).



(a) (b)

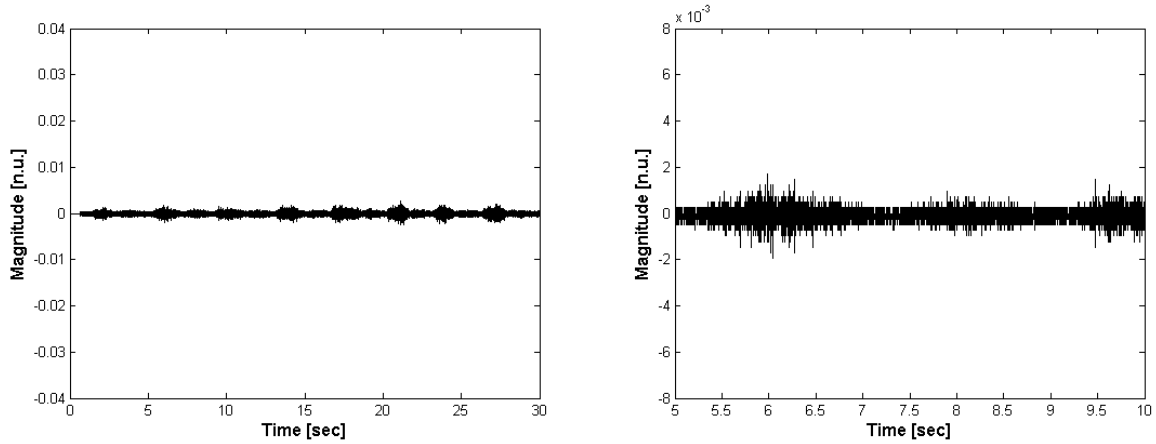
Figure 4-25. Uncompressed lung sound recording from healthy female adult using stethoscope device attached to Droid phone. Full 30-second signal (left) and five-second segment (right).

Recordings were performed under the same conditions except for the option to compress the signal when the Rehearsal Assistant program wrote the audio to a file. Otherwise, the test subject, recording locations on the body, stethoscope prototype, and Droid phone remained unchanged. The heart and lung sound waveforms derived from the compressed audio files (that were sampled at 8000 kHz) are illustrated in Figures 4-26 and 4-27. These figures are plotted on the same time and magnitude axes as Figures 4-24 and 4-25 for comparison. The compressed audio files are observed to have lower intensity waveforms than the uncompressed ones. It can also be deduced that the compression algorithm in the Rehearsal Assistant program filters out low frequency content in the critical range where heart and lung sounds occur. In this case, compression does not allow for faithful reproduction of the original signal. By focusing on five-second segments of the recordings, the loss of information can be seen in the blurry signals of Figures 4-26b and 4-27b, compared to the sharp signals in 4-24b and 4-25b.



(a) (b)

Figure 4-26. Compressed heart sound recording from healthy female adult using stethoscope device attached to Droid phone. Full 30-second signal (left) and five-second segment (right).



(a) (b)

Figure 4-27. Compressed heart sound recording from healthy female adult using stethoscope device attached to Droid phone. Full 30-second signal (left) and five-second segment (right).

To further demonstrate the consequences of audio file compression, the power spectral densities of the sounds with and without compression are plotted in Figure 4-28. The top two graphs represent uncompressed heart and lung sound recordings, respectively. Through visual inspection, the power at frequencies below 500 Hz is significantly lower in the compressed audio files compared to the uncompressed versions. In conclusion, this analysis suggests that a device's frequency and soundcard characteristics should be evaluated before use in a clinical setting, and that the transmission of data to the specialist should be verified as unaltered in order to reduce as many barriers as possible to enable accurate diagnosis of the patient.

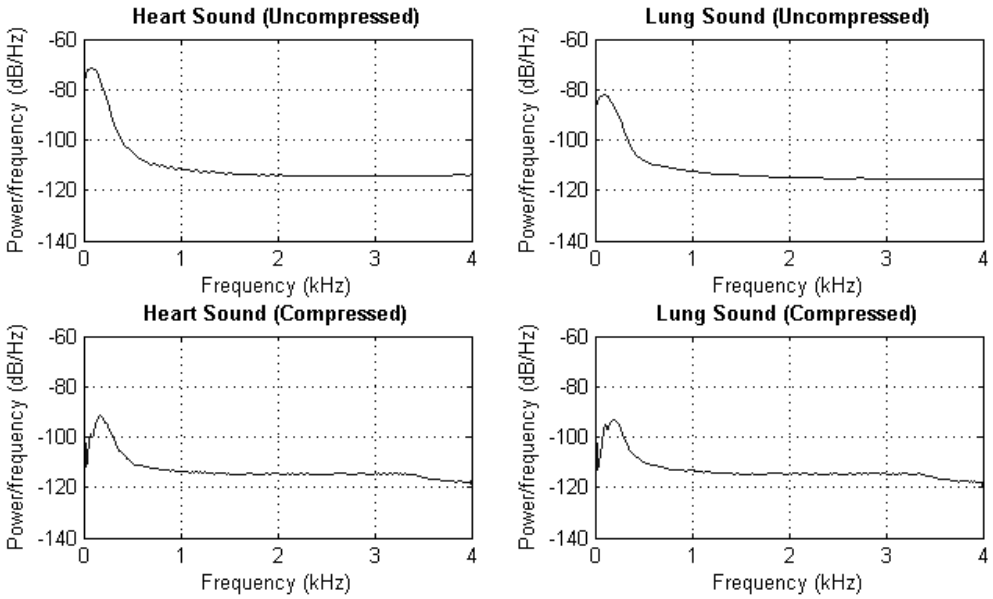


Figure 4-28. Power spectral densities of heart and lung sounds recorded with and without compression.

Chapter 5
Analysis of Field Data

5.1 Collecting Heart and Lung Sounds in Resource-Constrained Environments

In order to demonstrate the feasibility of capturing heart and lung sounds in resource-constrained environments (using technology with an existing supply chain infrastructure), audio data was collected with a mobile phone from 50 test subjects. The audio files were then transferred to a laptop for off-line analysis, using the heart and lung sound algorithms introduced in Chapters 2 and 3. The 400 mobile phone recordings constitute a more difficult set of algorithm test cases than the high-quality data used in earlier chapters (which were recorded with a commercial electronic stethoscope). This chapter discusses the protocol used for recording sounds from test subjects as well as the results from data analysis.

5.1.1 Equipment and Protocol for Recording Heart Sounds

Based on experience from two doctors and supporting evidence from Bates' guide to physical examinations (2005), the optimal locations on the chest for heart sound recording were selected, as depicted in Figure 5-1. Doctors with a traditional stethoscope typically only listen for several breaths or heartbeats at each location. However, mobile phone recordings introduce many variants of noise, so a 30-second

recording at each location should allow for some periods of inadvertent noise but also provide sufficient segments of noise-free audible data for the specialist to make an impression.

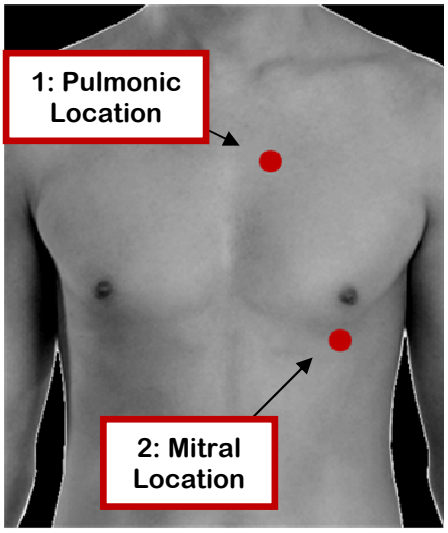


Figure 5-1. Two heart sound recording locations on a patient's front chest.
Adapted from Haggstrom (2009).

From Chapter 4, the Motorola Droid and Apple iPhone 3GS were found to be superior phones (in theory and in terms of technical specification). Since the Droid is an Android phone (on which the Sana telemedicine platform runs), this phone was selected for use in data collection with an accompanying hands-free kit and stethoscope device (see Figure 4-23). The Rehearsal Assistant Android application, which records uncompressed WAV files, was used to record a 30-second mono audio track from each subject. The patient's age, gender, weight, height, and relevant clinical history were recorded in a spreadsheet on a laptop.

In addition to having their heart sounds recorded, the subjects also had their heart rates simultaneously monitored with a $120 Contec CMS 50 fingertip pulse

oximeter[33] (Contec Medical Systems Co, Hebei, China). The pulse oximeter readings for each subject provided gold standard heart rate values[34] used in the evaluation of the accuracy of the acoustic heart rate estimation algorithm. The pulse oximeter (see Figure 5-2) was connected to a laptop for real-time transmission and display of pulse rate using the manufacturer's supplied software.

Figure 5-2. Contec pulse oximeter used for measuring heart rate in this work.

The protocol for recording heart sounds and pulse oximeter data was as follows:

1. The subject was seated in a chair and asked to sit upright. They were also asked to relax for three minutes while this procedure was explained to them. The subject was instructed to relax and breathe normally. They were also asked not to talk, change posture or move the finger, hand, or arm to which the pulse oximeter was connected. The subject was asked to lift their shirt, so that the rubber rim and plastic diaphragm of the stethoscope was in direct contact with their skin. For the first recording, the stethoscope was placed on the subject's pulmonic heart location (See Figure 5-1, position #1).

[33] A pulse oximeter measures arterial oxygen saturation by passing red and infrared light through an extremity and calculating the absorption (Greaves et al. 1997). The resulting pulsatile waveform can be used to calculate the subject's heart rate.

[34] The manufacturer did not provide specifics on the expected error rate of the pulse oximeter device, so it is possible that the reported heart rate estimation errors in this chapter may include the device's inherent measurement error.

2. Since it was not possible to listen to the heart sounds being recorded in real-time, an experimental 10-second recording was initially created. Then the recording was played to verify that heart sounds could be heard. If not audible, the home-made stethoscope was shifted a few centimeters in a clockwise radius from the original point of recording.
3. Once a suitable placement of the stethoscope was found, the pulse oximeter was attached to the subject's finger.
4. Once the pulse oximeter detected the subject's heart rate (evidenced by the display of a real-time waveform on the laptop monitor), heart sounds were recorded for 30 seconds at the pulmonic heart location. The recording was played afterwards to verify that the sounds were audible, otherwise the step was repeated.
5. The average heart rate[35] for the subject given by the pulse oximeter during the corresponding 30-second period was recorded on the computer, and was later associated with the correct audio recording.
6. Steps 4-7 were then repeated to capture heart sound data for 30 seconds at the mitral heart location.

5.1.2 Protocol for Recording Lung Sounds

In the same way that the heart sound recording locations were selected from expert consultation, the locations on the posterior chest selected for listening to lung

[35] The pulse oximeter software application did not allow access to the raw data, but did provide an average HR value in the recording's summary report. Details on how the average heart rate value was computed were not given by the manufacturer.

sounds were: the upper left, upper right, center left, center right, lower left, and lower right posterior chest (see Figure 5-3).

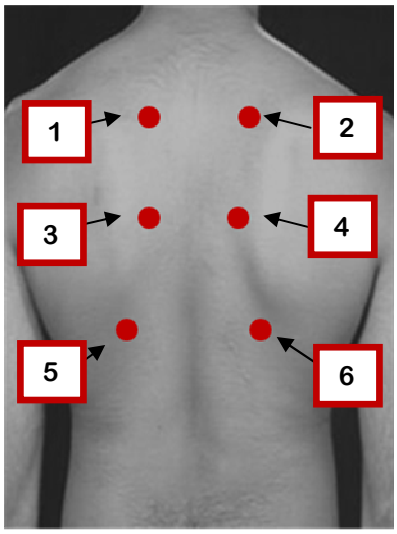


Figure 5-3. Six lung sound recording locations on a patient's posterior chest.
Adapted from Chandrasekhar (1999).

In order to determine a gold standard respiration rate for each lung sound recording, the study investigator manually counted and recorded the number of breath cycles completed by the test subject in the 30 seconds of the recording.

The lung sounds recording protocol was as follows:

1. The subject was positioned to have his or her back face the study investigator, while sitting upright. The subject was asked to lift his or her shirt when the stethoscope device was being positioned for recording.
2. The subject was instructed to breathe deeply through the mouth so that the study investigator could manually count respiratory breath cycles at the same time.
3. The stethoscope device was first placed on the subject's back at position #1 from Figure 5-3. Again, the rubber rim and plastic diaphragm of the device were held against the patient's back with direct skin contact and slight pressure.

4. Lung sounds were recorded first for 10 seconds to test signal quality. If the sounds were not audible, the device was shifted a few centimeters in a clockwise radius from the original point of recording. The stethoscope was not placed directly on the shoulder blades because the bone mass would muffle the lung sounds. However, if no other locations offered audible lung sounds, a partial section of the stethoscope diaphragm could be placed over areas with bone.
5. Once the location for an audible recording was found for position #1 in Figure 5-3, lung sounds were recorded for 30 seconds. The breath cycles were simultaneously counted by watching the rise and fall of the subject's shoulders. The resulting audio file was verified to have audible lung sounds (by listening for a sound similar to wind blowing through a tunnel).
6. The number of breaths completed during the corresponding 30-second period was recorded on the computer and associated with the correct audio file.
7. Steps 4-6 were repeated to capture lung sounds for 30 seconds at the five remaining locations in Figure 5-3.

5.1.3 Test Subjects

The demographics of the test subjects, including age, gender, height, weight, body mass index (described below), and clinical history, are given in Table 5-1. Each subject was allocated a unique numeric identification number (ID) and no personally-identifiable information was recorded to maintain the subjects' anonymity.

The height and weight of the subjects were collected in order to identify whether there were any correlations between algorithm performance and the size of the

individual. It is known that fat and muscle in the body can lead to decreased sound transmission because the sound must pass through dense tissue before reaching the surface of the body (McCann et al. 2002). Hence, body mass index (BMI) is one indicator used to quantify body fatness. An individual's BMI level can be classified under one of four categories: underweight (BMI of < 18.5), normal (18.5-24.9), overweight (25-29.9), or obese (≥ 30) (CDC 2009). The BMI was calculated for each individual using Equation 5.1, where W_{subj} represented the subject's weight in pounds and H_{subj} was the subject's height in inches (CDC 2009).

$$BMI = \frac{W_{subj} * 703}{(H_{subj})^2} \tag{5.1}$$

Among the 50 subjects, 8% were underweight, 72% were normal, 20% were overweight, and none were obese. It is also worthwhile to note that 20% of the study population was female, and only 14% were considered unhealthy.

ID	Age	Sex	Height (in.)	Weight (lbs.)	BMI	Clinical History
1	19	F	62	135	24.69	Healthy
2	23	M	73	185	24.41	Healthy
3	21	F	65	125	20.80	Healthy
4	23	M	68	145	22.04	Healthy
5	22	F	66	140	22.59	Healthy
6	25	M	72	160	21.70	Healthy, History of asthma
7	21	M	68	110	16.72	Healthy
8	44	M	65	160	26.62	Healthy
9	23	M	70	155	22.24	Unhealthy, Cough from a cold
10	22	F	59	120	24.23	Healthy
11	30	M	68	140	21.28	Healthy
12	22	M	67	150	23.49	Healthy
13	23	M	75	165	20.62	Healthy
14	21	F	62	110	20.12	Healthy
15	18	M	70	168	24.10	Healthy
16	18	M	72	170	23.05	Healthy
17	22	M	76	215	26.17	Unhealthy, Smoker for 4 years
18	21	M	70	170	24.39	Unhealthy, Smoker for 1.5 years
19	19	F	64	100	17.16	Unhealthy Smoker for 1 year
20	21	M	75	170	21.25	Healthy, Intoxicated
21	26	F	64	118	20.25	Healthy
22	22	M	66	150	24.21	Healthy
23	29	M	72	200	27.12	Healthy, Has earache
24	34	M	67	120	18.79	Healthy, Foot problem
25	22	M	69	145	21.41	Healthy
26	19	M	68	130	19.76	Healthy
27	23	M	68	132	20.07	Healthy
28	21	M	70	145	20.80	Healthy
29	25	M	71	154	21.48	Healthy
30	23	M	67	150	23.49	Healthy
31	23	M	69	155	22.89	Healthy, Had bronchitis before
32	23	F	66	110	17.75	Healthy
33	22	M	70	160	22.96	Healthy, Had asthma as a kid, had bronchitis before
34	27	M	68	165	25.09	Healthy, Asthma as a kid
35	32	M	73	160	21.11	Healthy
36	20	M	71	140	19.52	Healthy
37	19	M	66	150	24.21	Healthy
38	23	M	67	170	26.62	Healthy
39	32	M	74	195	25.03	Healthy
40	33	M	70	165	23.67	Healthy
41	23	M	68	165	25.09	Healthy
42	42	M	73	160	21.11	Sleep apnea, had bronchiolitis
43	32	M	71	185	25.80	Healthy, Had pneumonia
44	27	M	67	122	19.11	Healthy
45	27	M	72	185	25.09	Unhealthy, HOCM (hypertrophic obstructive cardiomyopathy) disease
46	23	M	65	122	20.30	Healthy
47	33	M	69	189	27.91	Unhealthy, Smoked for 11 years, mild asthma, worse when younger, had pneumonia, had collapsed lung
48	18	M	69	152	22.44	Healthy
49	22	F	64	105	18.02	Unhealthy, Cough from a cold
50	25	F	66	135	21.79	Healthy

Table 5-1. Basic demographic information on test subjects.

5.2 Heart Sounds Analysis

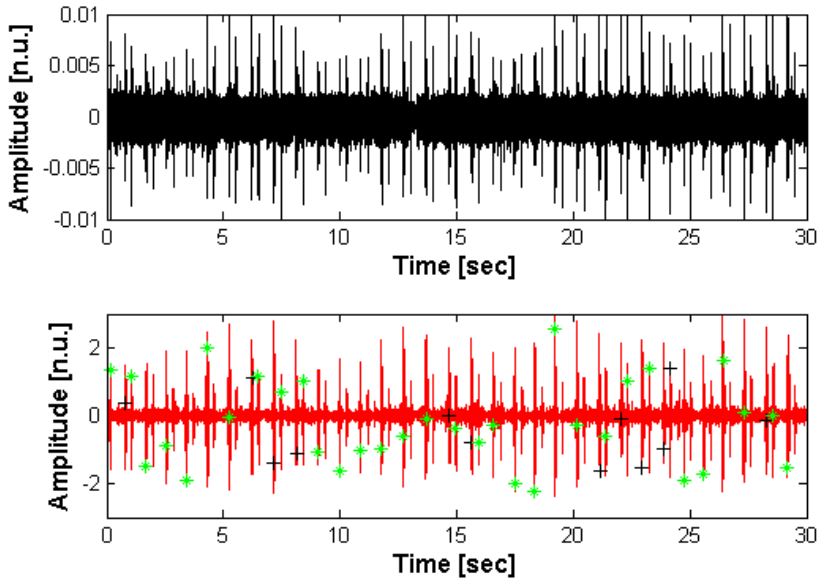
5.2.1 Original Heart Rate Estimation Algorithm

The acoustic heart sound method for heart rate estimation described in Chapter 2 was applied to the mobile phone recordings from test subjects. There were 100 audio files analyzed in total (two heart locations per subject) with the summary of results displayed in Table 5-2. The heart rate (HR) estimate from the algorithm was compared against the actual heart rate of the recording (provided by the pulse oximeter) to compute a percentage error for each record. The RMSE and RMSPE were subsequently computed for the dataset. If the algorithm was unable to distinguish S1 from S2 sounds based on the histogram of all SS intervals (see Section 2.2.3), no heart rate estimate could be provided (in 13% of the recordings), and these records were excluded from the final result statistics. A comprehensive table with the algorithm performance of each recording (actual HR, estimated HR, and percentage error) is provided in Appendix A. Each recording is labeled according to the subject ID number and recording location ID number (one or two).

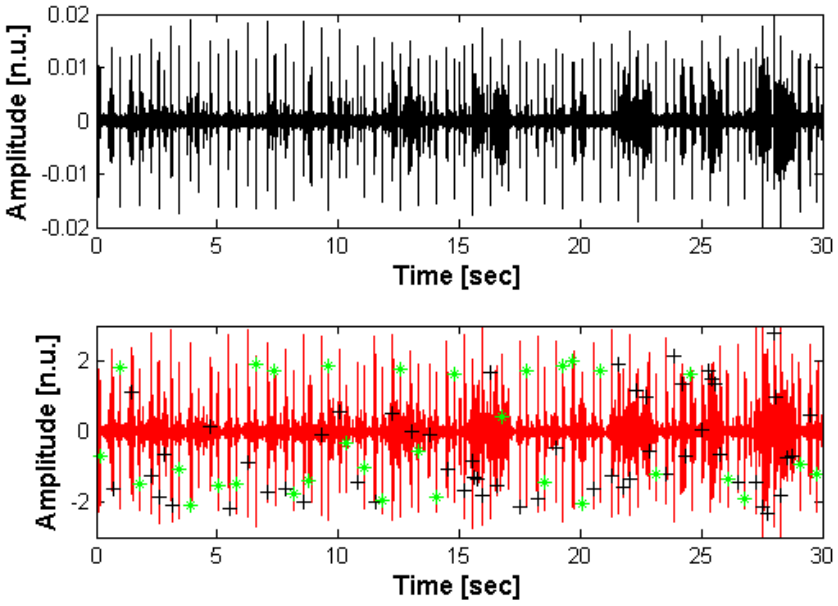
Results Summary	
Recordings Used	87%
Average Absolute Error	36.52%
Root Mean Squared Percentage Error (RMSPE)	6.30%
Root Mean Squared Error (RMSE)	1.09 bpm

Table 5-2. Summary of heart rate algorithm performance on test subjects.

Upon further investigation, the recordings (for which the algorithm was unable to produce HR estimates) were actually of sufficient quality such that heart beats could be manually counted from visual inspection. Figure 5-4 illustrates mitral heart sound recordings from subject 36 (part a) and subject 6 (part b).



(a)



(b)

Figure 5-4. Two example recordings of poor S1 and S2 detection from a) subject 36 and b) subject 6. In each part, the top plot shows the raw heart sound signal, while the bottom graph contains the processed signal with detected S1 (black '+' symbol) and S2 sounds (green '*' symbol).

From the lower plot in Figure 5-4a, it was observed that the algorithm skipped over the smaller peaks, because their amplitudes fell below the signal peak threshold. These peaks represented important S1 or S2 sounds, contributing to inaccurate SS interval calculations and inaccurate segmentation of S1 and S2 sounds. The inaccurate segmentation is also evidenced through the clusters of S1 sounds found in the waveform instead of the presence of alternating S1 and S2 sounds for consecutive peaks. Conversely, too many peaks were detected in other recordings (see lower plot of Figure 5-4b), especially in the presence of noise. This signal also had dense clusters of S1 sounds, which led to very small S1-S1 intervals for abnormally high HR estimates (based on Equation 2.11). For subject 6's recording in Figure 5-4a, the algorithm estimated a HR of 200 bpm while the actual HR was 79 bpm - a 153% percentage estimation error. As a result of the high error rates, modifications were developed for the algorithm, as described in Section 5.2.2. To reduce confusion, the original version of the algorithm will be referred to as HR_{est1}, while the updated one in Section 5.2.2 will be called HR_{est2}.

5.2.2 Modified Heart Rate Estimation Algorithm

The HR_{est2} algorithm was developed according to a training set of 50 recordings, which consisted of both mitral and pulmonic heart sounds from subjects 1-25 in Table 5-1. Among the subjects in the training set, 28% were females and 16% were considered unhealthy, the median age was 22, and the median BMI was 22.24.

To reduce the occurrence of clusters of peaks detected in noisy regions of a signal, a 200 ms refractory 'blanking' period was added into HR_{est1} after the integrated

waveform was generated. The refractory blanking rejected any detection less than 200 ms after a detected peak. While this feature was part of Pan and Tompkins' original QRS detector (Section 2.2.1), it was not previously included in the ECG and acoustic HR estimation algorithm used for heart sounds in Chapter 2.

The process of segmenting S1 and S2 sounds was described in step 5 of the HR_{est1} algorithm (in Section 2.2.3). The goal of this step was to separate the SS interval data in the histogram (see Figure 5-5) as belonging to one of two categories: S1-S2 or S2-S1 intervals. Slight modifications were added to this process in order to accommodate edge cases. The ability to detect two local maxima (or two separate clusters) in the data was added in the cases where peak values were equivalent or if there weren't local maxima[36] in the histogram data.

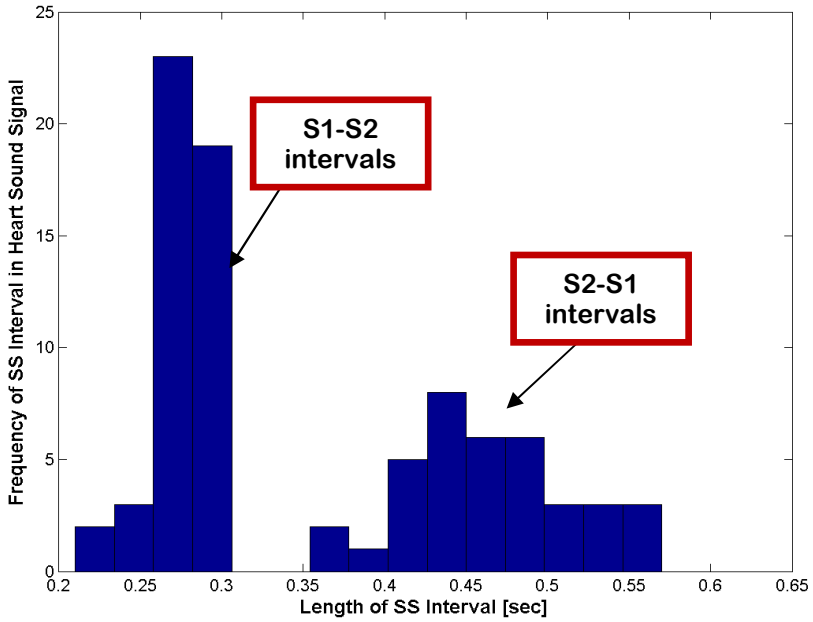


Figure 5-5. Example histogram with the frequency of SS intervals in a pulmonic heart sound recording from subject 39.

[36] Local maxima were defined as being higher than its two closest neighbors (on the immediate left and right) by at least a height of two.

Upper and lower adult heart rate limits were also integrated into the algorithm to recognize invalid estimates and reanalyze the data for a more accurate value. The normal heart rate for an average healthy adult at rest can range from 60 bpm to 100 bpm, per Dreeben (2008) and Fleming et al. (2010). From observation of the training set data[37] and to account for patients with diseases such as tachycardia, the lower limit of the acceptable HR range was reduced by a delta of 30 bpm and the upper limit was increased by 60 bpm (Dreeben 2008, Fleming et al. 2010). In this algorithm, the acceptable range of HR values was from 30 to 160 bpm. Note that this range was applicable for adults at rest, of which the study population was entirely composed. The range would have to be adjusted for children based on Fleming et al.'s (2010) work as described in Section 1.3.2.

When a HR estimate was detected outside the acceptable range, the algorithm dynamically adjusted its threshold level for peak detection (in a continuous loop) until a HR within the plausible range was discovered. Erroneous HR estimates in the range of 170 bpm to 400 bpm were observed with HR_{est1}, which meant that too many peaks had been detected and the S1-S1 intervals were too short. In order to reduce the number of peaks detected, the peak triggering threshold parameter, *t* (from Equation 2.10 in Section 2.2.2) was automatically raised by 0.01. Conversely, a HR estimate that was too low (below 30 bpm) implied that the S1 sounds were too far apart from each other and that not enough peaks were being detected. In that case, the threshold parameter *t* was decreased by 0.01 to increase the number of peaks detected. The value of *t* for each recording is given in the last column of the results in Table 5-3.

[37] Refer to Appendix B for the pulse oximeter heart rate readings for all field test subjects, who had a global HR minimum of 49 bpm and a global maximum of 99 bpm.

Another problem in heart beat detection was that the algorithm would label too many S1 sounds in a noisy segment of the signal (see bottom plot of Figure 5-6 at six seconds into the signal). In HR_{est1}, S1-S2 intervals corresponded to the cluster of shorter intervals (compared to S2-S1 intervals) in histograms similar to Figure 5-5. Sometimes when noise spikes in the data occurred closely together in a time series waveform and they were incorrectly detected as signal peaks, the noise spikes would be corresponded with short intervals in the histograms. The intervals between noise peaks consequently resided in the S1-S2 interval cluster (of the histogram), which led noise peaks to be incorrectly labeled as S1 sounds (see bottom plot in Figure 5-6).

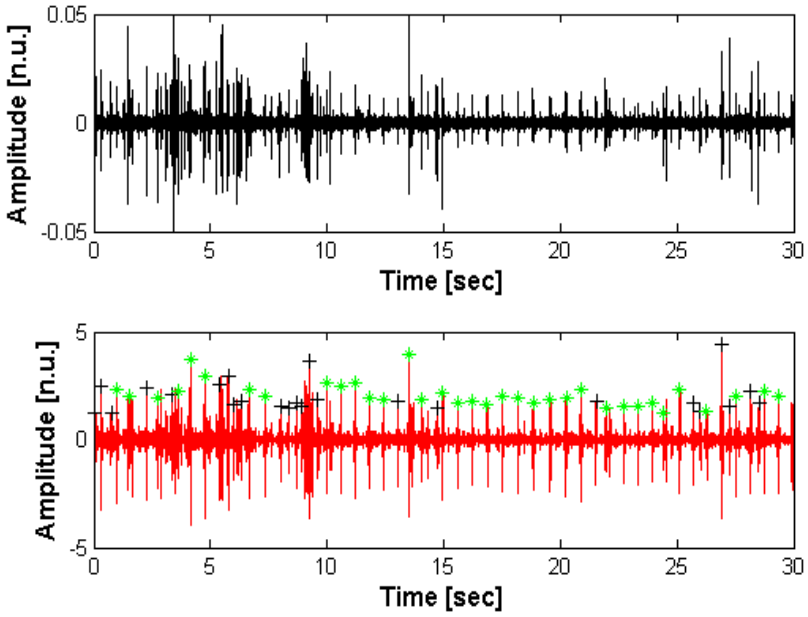


Figure 5-6. Recording with S2 sounds (green '*') are labeled more accurately than S1 sounds (black '+').
Top: Raw heart sound signal
Bottom: Processed signal from HR_{est1} algorithm with detected S1 and S2 sounds.

On the other hand, the cluster in the histogram corresponding to S2-S1 intervals typically did not include intervals between noise peaks. For each S2-S1 interval, the algorithm labeled S2 sounds on the processed signal. For example, in the lower plot of

Figure 5-6, S2 sounds were labeled for the majority of the heart beats in the signal, as seen by a human observer. Assuming that S2 sounds were labeled more accurately than S1 sounds in most of the field recordings, the HR could be calculated with Equation 2.11 as before, except with the S2-S2 intervals substituted for the S1-S1 intervals. Since each heart beat has one S1 and one S2 sound, it is possible to measure heart rate based on the time interval solely between S1 sounds or between S2 sounds without any difference in result. An example of the algorithm's improved performance with a HR estimate based on S2-S2 intervals can be seen in Figure 5-6 (from subject 18). Previously this subject's recording led to a 138% estimation error with HR_{est1}, but HR_{est2} produced an estimate with only 3% error.

The results of applying HR_{est2} on the training data set are shown in Table 5-3 with the location on the body of the recording, actual HR, estimated HR, percentage error, and threshold parameter value for triggering peaks. A performance summary is also provided in Table 5-4, where it can be seen that the RMSPE is 2.46% (compared to 6.3% previously) and 100% of the recordings are used.

SUBJ ID	LOCATION ID	ACTUAL BPM	ESTIMATED BPM	PERCENTAGE ERROR	THRESHOLD
1	**1**	71	72.82	0.03	0.08
1	**2**	74	75.57	0.02	0.08
2	**1**	89	86.71	-0.03	0.08
2	**2**	89	84.75	-0.05	0.08
3	**1**	74	72.29	-0.02	0.08
3	**2**	74	71.77	-0.03	0.08
4	**1**	77	73.71	-0.04	0.08
4	**2**	78	86.21	0.11	0.08
5	**1**	57	55.10	-0.03	0.08
5	**2**	59	57.58	-0.02	0.08
6	**1**	75	125.52	0.67	0.08
6	**2**	79	79.47	0.01	0.08
7	**1**	82	81.52	-0.01	0.08
7	**2**	84	82.19	-0.02	0.08

8	**1**	72	142.86	0.98	0.08
8	**2**	71	138.25	0.95	0.08
9	**1**	86	78.53	-0.09	0.08
9	**2**	83	76.73	-0.08	0.08
10	**1**	79	79.16	0.00	0.08
10	**2**	88	79.89	-0.09	0.08
11	**1**	74	65.22	-0.12	0.08
11	**2**	71	68.65	-0.03	0.08
12	**1**	78	78.74	0.01	0.08
12	**2**	75	74.35	-0.01	0.08
13	**1**	49	49.79	0.02	0.08
13	**2**	51	52.45	0.03	0.08
14	**1**	84	82.64	-0.02	0.08
14	**2**	88	87.46	-0.01	0.08
15	**1**	63	58.82	-0.07	0.08
15	**1**	89	71.09	-0.20	0.08
16	**1**	73	78.95	0.08	0.08
16	**2**	82	78.53	-0.04	0.08
17	**1**	91	90.50	-0.01	0.08
17	**2**	91	84.75	-0.07	0.08
18	**1**	93	91.74	-0.01	0.08
18	**2**	96	93.03	-0.03	0.08
19	**1**	90	87.46	-0.03	0.08
19	**2**	99	99.67	0.01	0.08
20	**1**	89	86.46	-0.03	0.08
20	**2**	94	63.42	-0.33	0.05
21	**1**	69	67.72	-0.02	0.08
21	**2**	67	68.65	0.02	0.08
22	**1**	80	61.10	-0.24	0.06
22	**2**	74	71.26	-0.04	0.08
23	**1**	75	113.64	0.52	0.08
23	**2**	70	65.79	-0.06	0.08
24	**1**	89	87.21	-0.02	0.08
24	**2**	91	91.88	0.01	0.08
25	**1**	49	55.56	0.13	0.08
25	**2**	59	38.14	-0.35	0.08

Table 5-3. Heart rate estimation results on training set of subject recordings.

Results Summary	
Recordings Used	100%
Average Absolute Error	11.64%
Root Mean Squared Percentage Error (RMSPE)	2.46%
Root Mean Squared Error (RMSE)	2.95 bpm

Table 5-4. Summary of heart rate algorithm performance on training set of subjects.

5.2.3 Results for the Modified Heart Rate Estimation Algorithm on Test Data

The HR_{est2} algorithm was then applied to the test set of data containing recordings from the remaining 25 subjects (that were not included in the training set). The test set consisted of subjects 26-50 (from Table 5-1), where 12% were female, 12% were unhealthy, the median age was 23, and the median BMI was 22.44. Table 5-5 provides a detailed list of the algorithm results on each recording in the test set, while Table 5-6 offers a comparison of the performance between the training and test sets.

SUBJ ID	LOCATION ID	ACTUAL BPM	ESTIMATED BPM	PERCENTAGE ERROR	THRESHOLD
26	**1**	82	87.46	0.07	0.08
26	**2**	85	84.27	-0.01	0.08
27	**1**	82	81.86	0	0.08
27	**2**	82	81.97	0	0.08
28	**1**	66	72.21	0.09	0.08
28	**2**	80	77.72	-0.03	0.08
29	**1**	69	57.03	-0.17	0.08
29	**2**	62	86.71	0.4	0.08
30	**1**	53	52.82	0	0.08
30	**2**	59	56.5	-0.04	0.08
31	**1**	93	86.46	-0.07	0.08
31	**2**	84	80	-0.05	0.08
32	**1**	83	80.65	-0.03	0.08
32	**2**	94	82.87	-0.12	0.08
33	**1**	55	79.06	0.44	0.08
33	**1**	61	71.6	0.17	0.08
34	**1**	61	59.94	-0.02	0.08
34	**2**	68	66.67	-0.02	0.08
35	**1**	49	47.85	-0.02	0.08
35	**2**	58	52.63	-0.09	0.08
36	**1**	83	67.26	-0.19	0.08
36	**2**	72	64.79	-0.1	0.08
37	**1**	88	47.77	-0.46	0.08
37	**2**	68	47.24	-0.31	0.08
38	**1**	72	71.6	-0.01	0.08
38	**2**	74	72.82	-0.02	0.08

39	**1**	69	58.82	-0.15	0.08
39	**2**	60	59.76	0	0.08
40	**1**	79	74.44	-0.06	0.08
40	**2**	71	68.34	-0.04	0.08
41	**1**	78	70.59	-0.1	0.08
41	**2**	85	71.77	-0.16	0.08
42	**1**	81	83.33	0.03	0.08
42	**2**	88	84.87	-0.04	0.08
43	**1**	57	52.13	-0.09	0.08
43	**2**	56	54.95	-0.02	0.08
44	**1**	71	64.38	-0.09	0.08
44	**2**	63	58.94	-0.06	0.08
45	**1**	77	64.38	-0.16	0.08
45	**2**	73	124.74	0.71	0.08
46	**1**	61	51.64	-0.15	0.08
46	**2**	58	56.13	-0.03	0.08
47	**1**	93	59.45	-0.36	0.08
47	**2**	99	125.52	0.27	0.08
48	**1**	67	65.79	-0.02	0.08
48	**2**	68	48.78	-0.28	0.01
49	**1**	79	65.08	-0.18	0.08
49	**2**	67	63.03	-0.06	0.08
50	**1**	79	77.32	-0.02	0.08
50	**2**	68	75.38	0.11	0.08

Table 5-5. Heart rate estimation results with HR_{est2} on test set of subject recordings.

	Training Set	**Test Set**
Recordings Used	100%	100%
Average Absolute Error	11.64%	12.19%
Root Mean Squared Percentage Error (RMSPE)	2.46%	1.88%
Root Mean Squared Error (RMSE)	2.95 bpm	1.04 bpm

Table 5-6. Summary of HR_{est2} heart rate algorithm performance on training and test sets.

The performance indicators in Table 5-2 and Table 5-6 demonstrate that the HR_{est2} algorithm showed improved estimation accuracy on the subject recordings over HR_{est1}. The 1.88% RMSPE and 1.04 bpm RMSE of the HR_{est2} method on the test set were much better than the HR_{est1} method's 6.3% RMSPE and 1.09 bpm RMSE error levels.

An example of accurate heart beat detection on a pulmonic heart sound recording from subject 14 is shown in Figure 5-7. Although the raw signal contained some noise, the processing of the signal produced clear enough peaks for a human to be able to manually compute the heart rate from the S1 and S2 sounds. For each beat, an S1 and S2 sound were both identified by HR_{est2} in an alternating fashion, so the resulting HR estimate could have been computed either from S1-S1 intervals or S2-S2 intervals. The HR_{est2} method estimated a HR of 82.6 bpm (compared to the actual 84 bpm) - a 2% absolute estimation error.

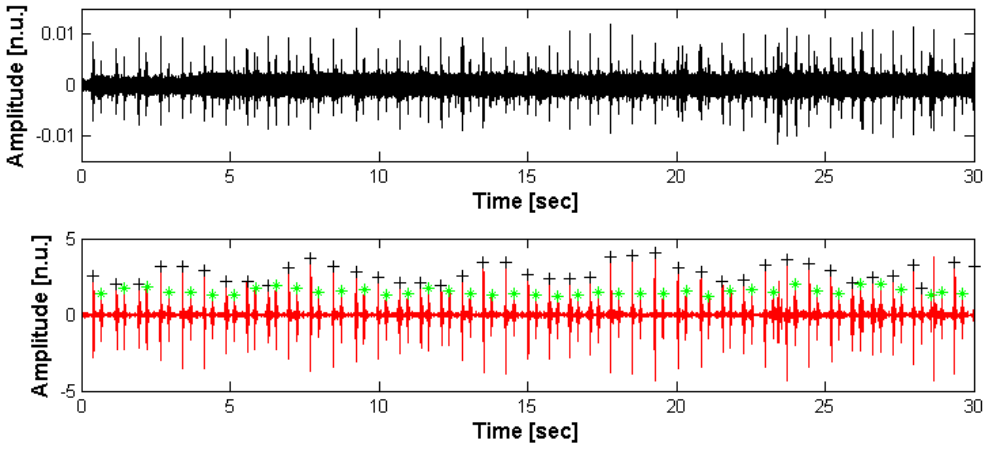


Figure 5-7. Example recording with accurate S1 and S2 detection.
Top: Raw heart sound signal
Bottom: Processed signal from HR_{est2} algorithm with detected S1 ('+') and S2 sounds ('*').

Poor heart beat detection performance is shown on a pulmonic recording from subject 4 in Figure 5-8. From examination of the raw signal, there were numerous large noise peaks that had amplitudes many times larger than the S1 and S2 peaks (see the signal between 23 to 26 seconds in the upper plot of Figure 5-8). The algorithm unsuccessfully labeled the signal with S1 and S2 sounds, although a human annotator would likely also have trouble locating the S1 and S2 sounds.

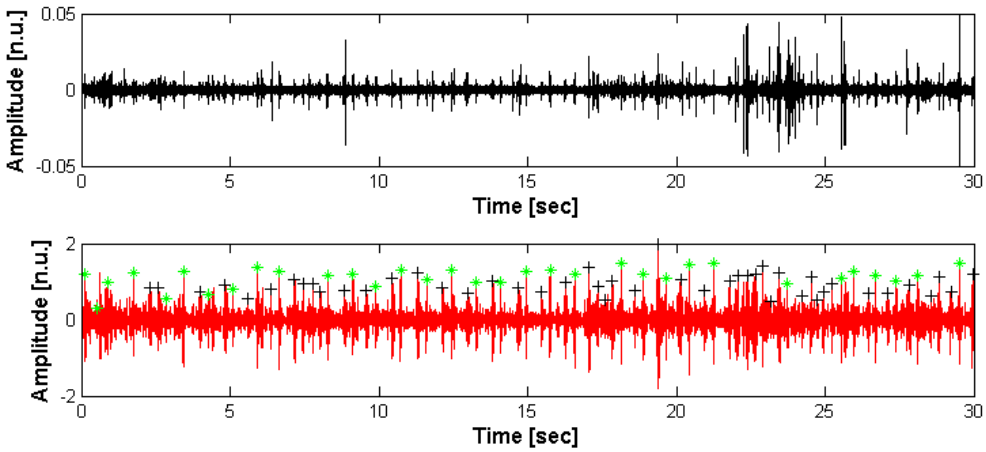


Figure 5-8. Example recording of healthy test subject with poor S1 and S2 detection.
Top: Raw heart sound signal
Bottom: Processed signal from HR_{est2} algorithm with detected S1 ('+') and S2 sounds ('*').

For a signal with reasonably good quality (visible heart beats), the presence of a large noise spike had the ability to significantly degrade the quality of heart beat detection over the whole recording, as seen in Figure 5-9. If high frequency noise was not removed by the band pass filter between 5 to 70 Hz in HR_{est2}, then a noise spike elevated the peak detection threshold P_{thresh} (in Equation 2.10) such that real heart sounds were missed. The HR_{est2} algorithm compensated for this problem by iteratively lowering the threshold parameter t until $t = 0.05$ (so P_{thresh} would be 5% of the max peak height in signal s_{int} from Equation 2.10). For the mitral heart sound recording of subject 20 in Figure 5-9, the HR_{est2} method produced an estimate of 63.42 bpm within the acceptable HR range. However, there were still many missed peaks because the actual HR was 94 bpm, resulting in a 33% absolute estimation error.

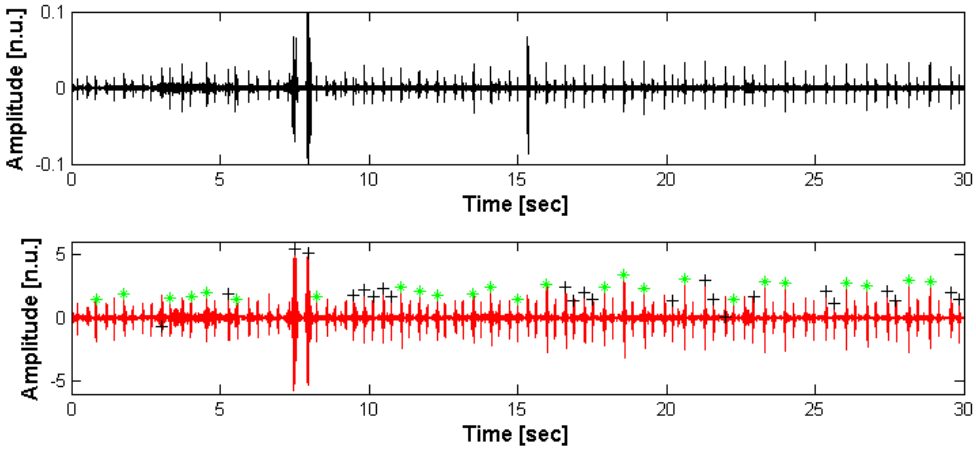


Figure 5-9. Example recording with noise peaks for poor S1 and S2 detection.
Top: Raw heart sound signal
Bottom: Processed signal from HR_{est2} algorithm with detected S1 ('+') and S2 sounds ('*').

The study population contained one individual with hypertrophic obstructive cardiomyopathy (HCOM) (see Figure 5-10). This type of heart disease is characterized by inflammation of the heart muscle in the left ventricle, where the septum separating the left and right ventricles is enlarged (American Heart Association 2010). Blood flow is obstructed, leading to symptoms such as dizziness, chest pain, murmurs, or heart arrhythmias with a risk of sudden death. In the heart sounds of HCOM patients, the normal clarity of the 'lub dub' sound of the heart beat is reduced. In addition, an S4 heart sound can be present or the S2 sound can be split (Douglas et al. 2009). The interference of additional sounds to the regular S1 and S2 sounds was seen in the mitral heart sounds of subject 47. In Figure 2-10, the signal peaks of S1 and S2 sounds were not easily distinguishable for subject 47. The HR_{est2} algorithm estimated a HR of 125.5 bpm instead of an actual 94 bpm, leading to a 27% absolute error. It is difficult to draw any general conclusions from two recordings of a heart disease patient because the waveform could have been contaminated with external ambient noise rather than abnormal heart sounds.

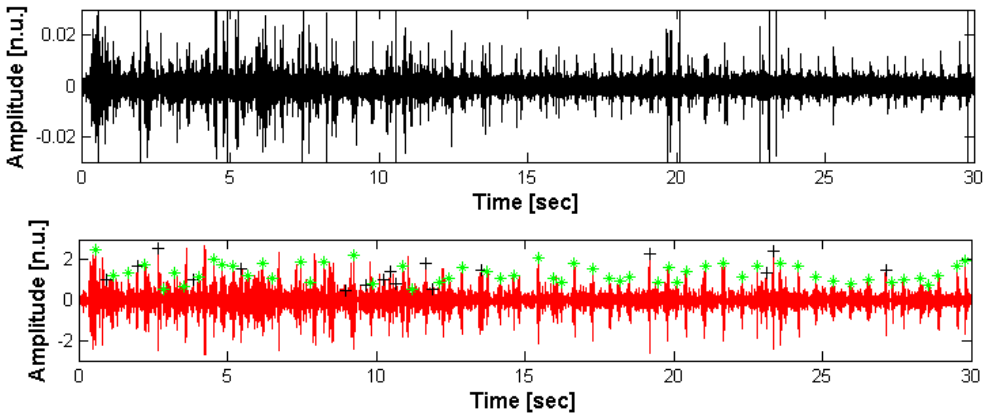


Figure 5-10. Test subject with abnormal heart condition (hypertrophic obstructive cardiomyopathy).
Top: Raw heart sound signal
Bottom: Processed signal from HR_{est2} algorithm with detected S1 ('+') and S2 sounds ('*').

As a broader analysis of the whole dataset (both test and training sets), Figure 5-11 displays the relationship between the actual HR of subjects in the recordings and the absolute percentage estimation errors (from the HR_{est2} method). A least squares fit was found to be approximately flat with $R^2 = 0.002$.

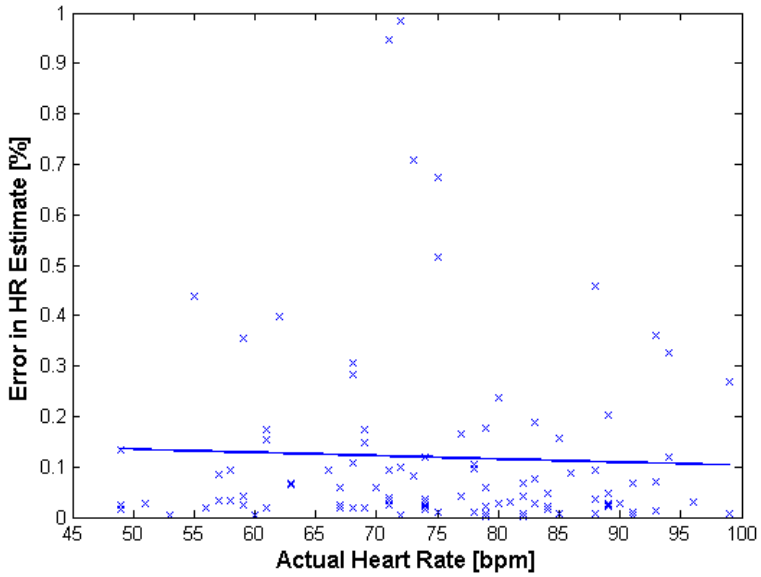


Figure 5-11. Actual Heart Rate (bpm) versus Absolute % Estimation Error with HR_{est2},
Best fit line y=0.167-(6.37e-4)*x with R^2=0.002

Another noticeable trend is illustrated in Figure 5-12, which shows the absolute percentage estimation error increased at higher BMIs.

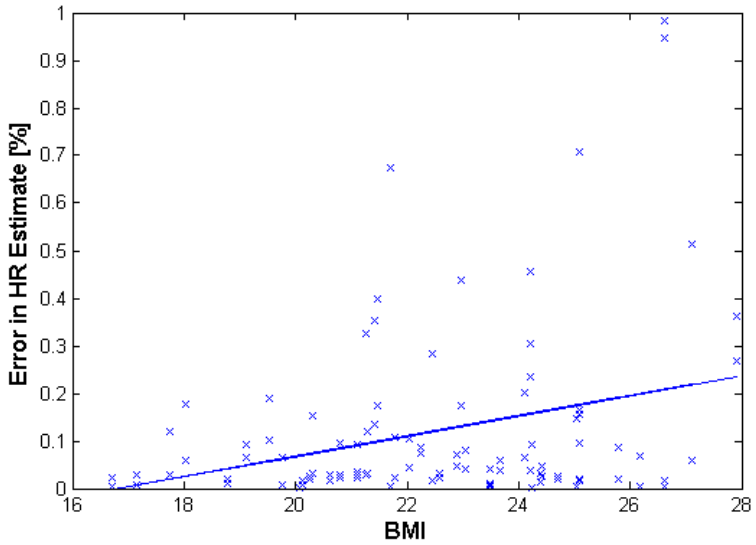


Figure 5-12. Actual Heart Rate (bpm) versus Absolute % Estimation Error, Best fit line y=-0.358+0.021x with R^2=0.094

Since the audio files were labeled according to location of the recording device, the data were also separated into mitral and pulmonic recordings. Figure 5-13 illustrates the dependence of the estimation error on HR for the pulmonic (left plot) and mitral regions (right plot). No large differences in performance were detected, since the best linear fit appeared to be similar (approximately flat) for both locations.

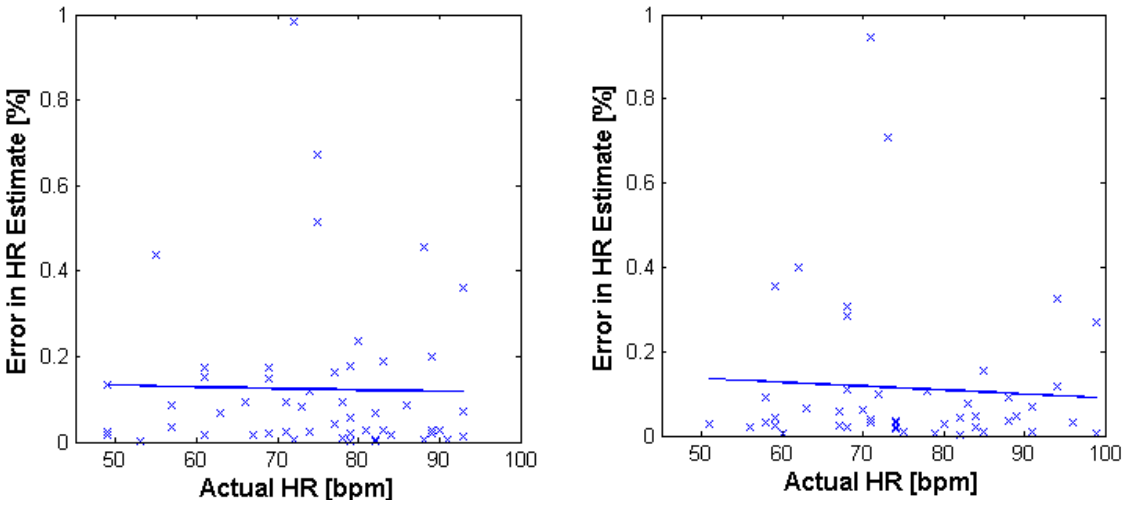


Figure 5-13. Comparison between recordings at pulmonic heart location versus mitral location
(left) Pulmonic: Actual heart rate versus absolute % estimation error,
Best fit line y=0.15-(3.53e-4)*x with R^2=5.37e-4
(right) Mitral: Actual heart rate versus absolute % estimation error,
Best fit line y=0.184-(-9.34*e-4)*x with R^2=0.004

5.2.4 Discussion of Heart Rate Estimation Results

Since mobile phones are portable and allow for heart sounds to be recorded in any public environment, there were varying levels of noise that made algorithm parameters suitable for one recording inaccurate for another. The HR_{est2} method was shown to have better detection capabilities than HR_{est1} because of the HR_{est2} method's dynamic ability to adjust the peak triggering threshold value to produce an optimal estimate within a plausible range of heart rates.

However, in the example recordings where heart beats could not have been clearly labeled by a human observer due to the interference of noise peaks (Figures 5-8 and 5-10), the HR_{est2} algorithm produced poor heart beat detection. This challenge emphasizes the importance of having a signal quality check before performing analysis. The signal purity measure based on the sinusoidal nature of a lung sound recording may not work as well on heart sound recordings. Different heart sounds (S1, S2, S3,

and S4) have varying intensities (where S3 and S4 sounds are sometimes not audible at all), so a noise-free heart sound waveform would not be perfectly sinusoidal with a fixed amplitude.

The ability to detect the presence of noisy segments in a signal could potentially aid in the development of signal quality measures and could also be leveraged to reduce error in S1 and S2 detection. For example, the labeling of many consecutive S1 sounds in a noisy region of the signal (to indicate many S1-S2 intervals) is one sign that the detection algorithm is performing incorrectly. This behavior is partially attributed to the use of a histogram for identification of S1 and S2 sounds. The success of the histogram method is dependent on the number of buckets used in the histogram and the ability to cleanly split the data into two clusters (with a single peak in the center of each cluster). In the future, other clustering methods should be attempted on the histogram or alternative methods should be explored to ensure that each S1-S2 interval in the signal is followed by a S2-S1 interval (instead of being followed by another S1-S2 interval).

From observation of the test and training set data together (see Figure 5-12), a trend of increased absolute estimation error rate with increased BMI levels can be suggested. This relationship is consistent with the expectation that heart sounds are less audible on patients with a lot of body fat or muscle, leading to audio recordings with lower signal quality, and hence decreased algorithm detection capabilities, and higher error rates.

Figure 5-13 suggests that the HR_{est2} algorithm's estimation abilities are similar for both mitral and pulmonic heart sound recordings. This is an interesting observation

because S_1 sounds are heard loudest at the apex of the heart near the mitral location, while S_2 sounds are louder than S_1 sounds at the pulmonic location (McCann et. al. 2005a). It is worth exploring whether the larger peaks in a mitral heart sound recording can be classified as S_1 sounds and the reverse for larger peaks in pulmonic recordings with S_2 sounds. Further work must be also be done to confirm whether the S1 and S2 sounds were labeled properly by the HR_{est2} algorithm in this analysis, or whether the labels should be reversed (i.e. a detected S1 sound is actually an S2 sound). Afterwards, differences between mitral and pulmonic heart sound audio waveforms can potentially be linked to physiological causes.

5.3 Lung Sound Analysis

5.3.1 Lung Sound Recordings

In a similar way that heart sound recordings were analyzed, the respiratory sound analysis algorithms described in Chapter 3 were applied to the lung sound recordings from the study population. There were six lung sound recordings per patient, meaning 300 lung sound audio files in total. Each recording was labeled according to the subject ID number (1 to 50) and recording location ID number from (1 through 6). For each recording, the algorithms' respiratory rate estimates were compared with the gold standard respiratory rate value calculated from the manually-counted breath cycles.

The algorithms from Chapter 3.2 are described here again briefly. The Tracheal Breath Count (TBC) method obtains the envelope of the lung sound signal (through

filtering and downsampling), computes the derivative of the result, and then performs peak detection to determine the number of respiratory phases in the signal. The Modified Pan and Tompkins' (MPT) method uses the lung sound signal envelope (from the TBC method), integrates the envelope signal, and then performs peak detection on the result with a specific triggering threshold. The Autoregressive Model 1 (ARM1) method fits an AR model to the lung sound signal envelope (from the TBC method), and identifies the lowest angle pole (in the z-plane) that corresponds to a respiratory rate within the range of physiologically plausible rates. The ARM2 method is similar to ARM1 except that the ARM2 approach fits an AR model to the integrated waveform from the MPT method. The ARM3 method is also similar to ARM1 except that in the ARM3 approach, only the poles with magnitude ≥95% of the highest magnitude pole in the range of physiological plausibility are kept as candidates, from which the lowest angle pole is selected as the respiratory rate estimate. Finally, the MEDIAN method uses the median of the estimates generated from the TBC, MPT, and ARM1 methods.

5.3.2 Results on Lung Sound Recordings

A summary of the results is presented in Table 5-7. (See Appendix C for full details of the results from the 300 recordings).

Method	Recordings Used [%]	Average Absolute Error [%]	Root Mean Squared Error [%]	Root Mean Squared Error [cpm]
TBC	100	78.17	10.56	13.24
MPT	100	88.49	10.91	14.69
ARM1	100	87.31	8.90	14.70
ARM2	100	86.29	8.64	14.68
ARM3	100	87.31	8.90	14.70
MEDIAN	100	65.29	8.38	11.39

Table 5-7. Summary of respiration rate algorithm performance on 50 test subjects using the mobile phone and home-made stethoscope.

Based on the outcomes listed in Table 5-7, the MEDIAN method gave the lowest error with a RMSPE of 8.38% compared to the ARM2 method with 8.64%, ARM1 method with 8.90%, and MPT method with 10.91%. In addition, the RMSE was 11.39 cpm for the MEDIAN method compared to 14.68 cpm for the ARM2 method, 14.70 cpm for the ARM1 method, and 14.69 cpm for the MPT method.

A significant observation was that the TBC method (the best method on the lung sound database in Chapter 3) performed the second poorest among all the estimators for the field data with an RMSPE of 10.56% and RMSE of 13.24 cpm.

Despite the similar approach in the ARM methods, the ARM2 method had slightly better performance than the ARM1 and ARM3 ones (the latter two methods produced the same respiration rate estimates and error rates). The ARM2 method had a lower RMSPE by 0.26%, lower RMSE by 0.02 cpm, and lower average absolute error by 1.02% compared to the ARM1 and ARM3 methods.

5.3.3 Discussion of Respiration Estimation Results

The most important observation of the above results is that the MEDIAN method, and not the TBC method, was the best overall estimator. These results demonstrate that a 'mixture of experts' algorithm is able to out-perform any single algorithm because of the differences in the noise sensitivity profiles of each algorithm. Alternative more sophisticated data fusion frameworks (which factor in signal quality and previous history) may prove even more superior than any single method and this simple MEDIAN approach (Clifford et al. 2009).

The TBC method's high error rates may be attributed to the fact that the algorithm was developed for tracheal sounds and all the respiratory sounds collected from the subjects were from the lungs. The large and distinct energy bursts per breath cycle in tracheal sounds were not as pronounced in lung sounds, as shown in Figure 5-14. The TBC method depended on clear peaks and troughs in the signal, while the respiratory cycles in the mobile phone recordings sometimes could not be distinguished from one another through visual inspection. Furthermore, the superior recording conditions (protocol and/or equipment) of the gold standard database to capture subtle respiratory sounds evidently do not reflect the conditions that can be observed in field recordings.

Although the lung sound recordings were noisy, the ARM methods seemed to perform reasonably because the energy in the recordings followed a general oscillatory pattern indicative of respiration. The ARM2 method had higher accuracy than the ARM1 and ARM3 methods potentially because energy and integration-based algorithms are known for being robust methods to reduce sudden transients, or 'spikes' (Friesen et al. 1980). Although the MPT method also uses the integration step found in the ARM2 method, the MPT algorithm is sensitive to high energy transients, because the triggering threshold can be made too high in these scenarios. Moreover, the MPT method requires clear signal peaks in order to find each breath cycle, which were not apparent in the subject recordings and led to high error rates.

The waveforms of the audio files from the test subjects exhibited a wide range of perceived signal quality.[38] This can be attributed to either the diverse recording

[38] Signal quality in this context refers to the ability to qualitatively observe the respiratory cycles in the waveform.

environments or nature of different people's lung sounds. Three signals of varying quality are shown in Figure 5-14. The recording at the top of Figure 5-14 shows a "good" signal from the upper left back location of subject 25. Eight respiratory cycles can be seen in the 30-second waveform for a respiration rate of 16 cpm, which agrees with the value obtained from manually counting the subject's breaths at the time of the recording. The raw lung sound signal pictured in the center of Figure 5-14 was from location 5 of subject 14 (with an actual respiratory rate of 15 cpm), while the bottom plot in Figure 5-14 was from location 3 of subject 22 with a 14 cpm respiratory rate. The latter two plots have large noise disturbances, making the manual observation of distinct breath cycles difficult.

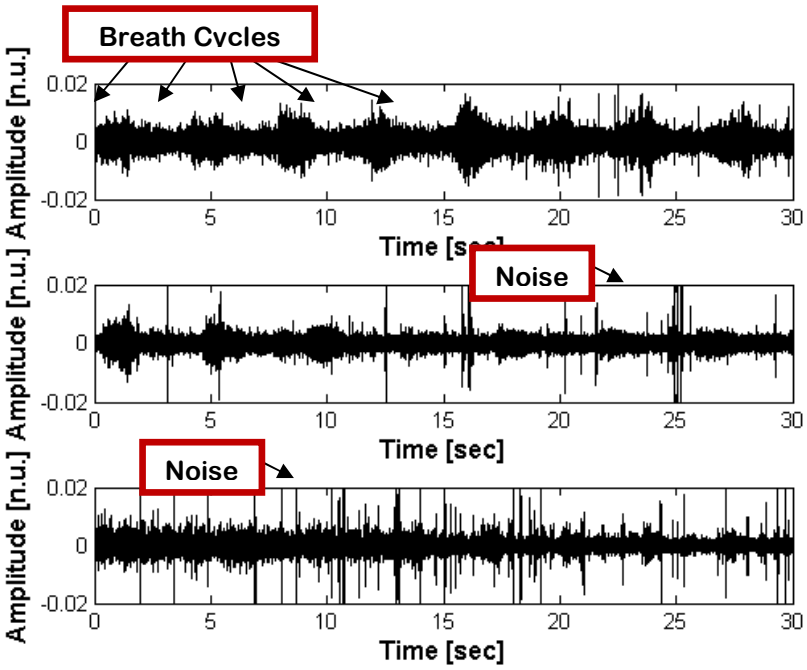


Figure 5-14. Lung sound recordings of varying quality from good (top) to medium (middle) and bad (bottom).

Several unhealthy subjects were also recorded to determine whether mobile phones would be able to capture heart and lung sounds as effectively as for healthy individuals. A typical lung sound waveform of a subject who had been smoking for four years is given in Figure 5-15 (subject 17, locations 4, 5, and 6). It appeared that his lung sounds were much quieter and had more irregular signal artifacts than in normal healthy patients. However, there was still concern that the artifacts were due to noise with the recording equipment instead of clinical indicators, so more smokers should be tested. There were also two other individuals in the study population who had started smoking in the past two years, but their lung sounds appeared and sounded normal. It is possible that the effects of smoking may not have had significant damage on their lungs by the time of recording. However, the small number of subjects in this category makes it hard to draw any conclusions with confidence.

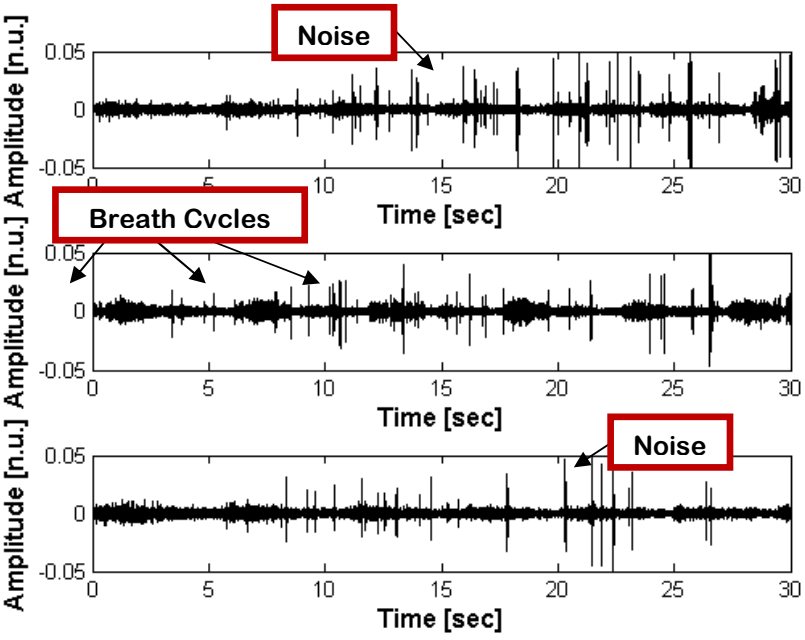


Figure 5-15. Lung sound recordings from 4-year smoker test subject.
From center right back (top), lower left back (middle), and lower right back (bottom).

Another unhealthy subject had asthma and was also a smoker for 11 years. Three of his lung sound recordings with audible wheezing are illustrated in Figure 5-16. The recordings were taken at locations 2, 3, and 4 on the subject's back with actual (observed) respiratory rates 17 cpm, 19 cpm, and 17 cpm respectively. Because wheezes are louder than normal vesicular lung sounds (see Section 1.2.3), they are picked up more easily and clearly by the stethoscope device. Wheezes are also musical in quality, which are reflected by the continuous tapering-off effect in the amplitude of each energy burst (McCann et al. 2002).

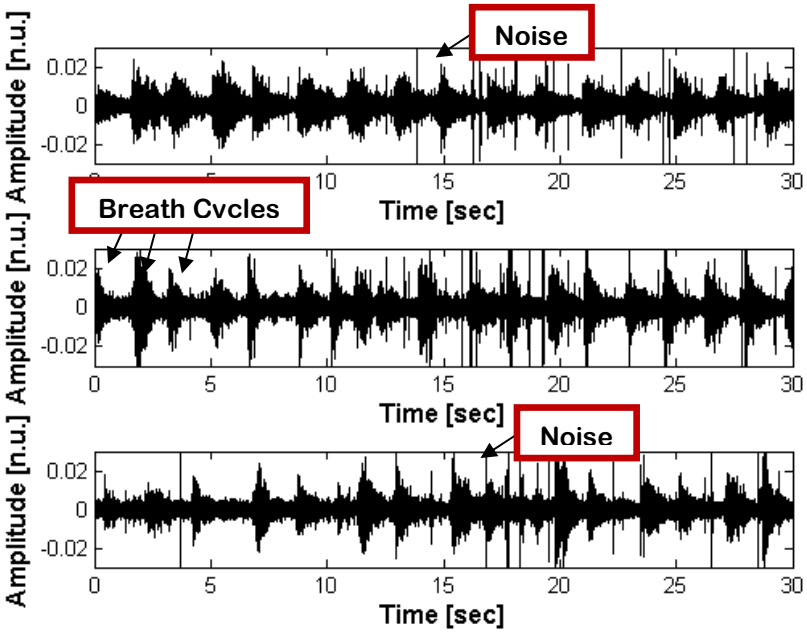


Figure 5-16. Lung sound recordings from a wheezing asthma test subject.
From upper right back (top), center left back (middle), and center right back (bottom).

Another class of respiratory sounds that would be interesting to explore in the future is coughing. Two subjects had a cough resulting from the common cold, which were captured in eight-second recordings seen in Figure 5-17. The plots in the left column present the full waveforms for subject 9 (top) and subject 49 (bottom). The plots

in the right column show 1.5-second segments of the recordings where the cough actually occurred. From Figure 5-17, it can be deduced that each cough produced intermittent large bursts of energy within the respiratory waveform (with amplitudes up to 40 times greater than those of normal lung sound recordings from Figure 5-14). Cough recordings can be useful in determining the frequency and severity of symptoms in a disease such as chronic cough (Krajnik et al. 2001).

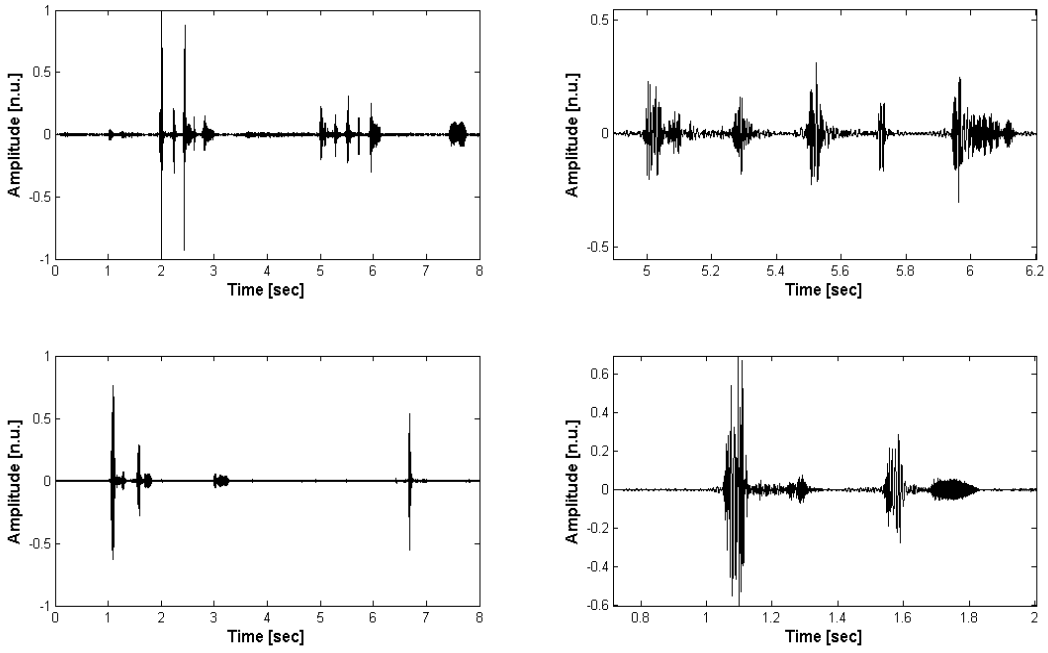


Figure 5-17. Cough recordings.
Adult male: full waveform (top left) and 1.5 second segment (top right)
Adult female: full waveform (bottom left) and 1.5 second segment (bottom right).

In Figure 5-18, the influence of actual respiratory rate on estimation error for a method can be observed for the lung sound recordings from all 50 subjects. The TBC, MPT, and MEDIAN method plots suggested that as actual respiratory rate increased, absolute percentage estimation error decreased. It was possible that when the subjects with high respiratory rate were breathing rapidly, they were breathing more forcefully,

leading to higher intensity waveforms that were more suitable for the TBC, MPT, and MEDIAN algorithms. The fact that the MEDIAN estimator showed decreased error for high respiratory rates is a promising indication that it can work as effectively on mobile phone recordings from subjects under 18 years old, who have normal respiratory rates above 20 cpm (Fleming et al. 2010).

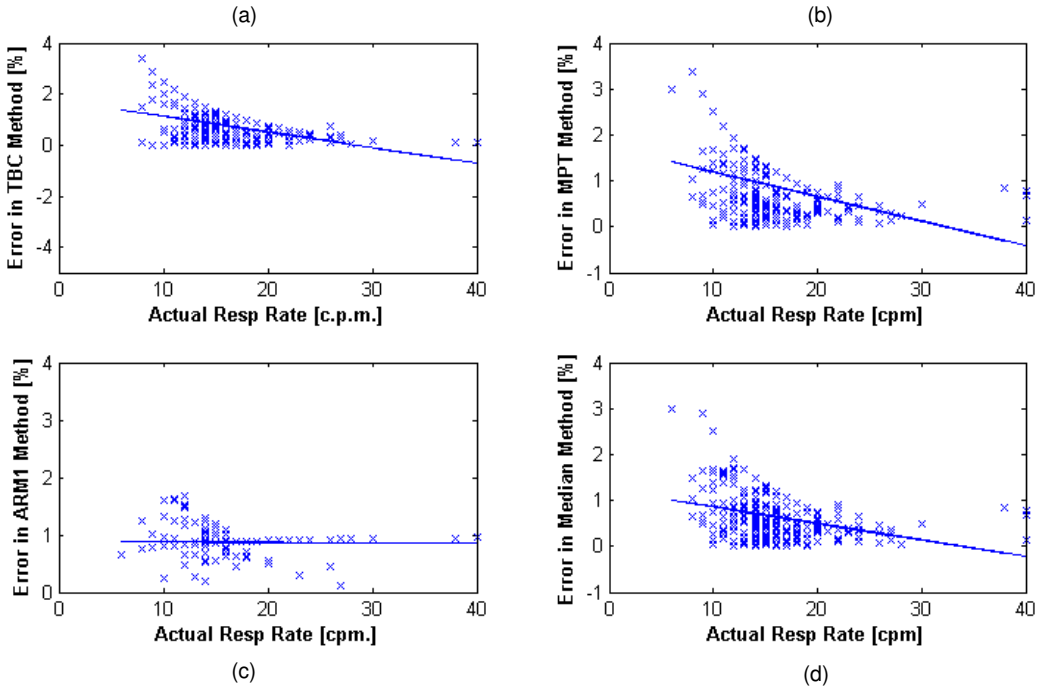


Figure 5-18. Absolute % error vs. respiration rate in test subject lung sound recordings.
a) TBC method, best fit line y=1.760-0.062x with R^2=0.193
b) MPT method, best fit line y=1.731-0.054x with R^2=0.179
c) ARM1 method, best fit line y=0.888-(9.17e-4)*x with R^2=7.06e-4
d) MEDIAN method, best fit line y=1.227-0.036x with R^2=0.121

Signal quality was discussed previously in a qualitative manner for Figure 5-16. The quality of lung sound recordings was also numerically quantified with a signal purity index based on Hjorth descriptors, as described in Section 3.2.5 (Nemati et al. 2010). Purity may range from 0 (poor) to 1 (excellent) and was measured on the envelope of filtered lung sound signals. For the entire dataset of lung sound mobile phone recordings, the purity levels ranged from 0.01 to 0.21, with an average purity of 0.05.

Figure 5-19 plots the best fit line for the relationship between purity level and absolute percentage estimation error with the TBC, MPT, ARM1, and MEDIAN methods separately.

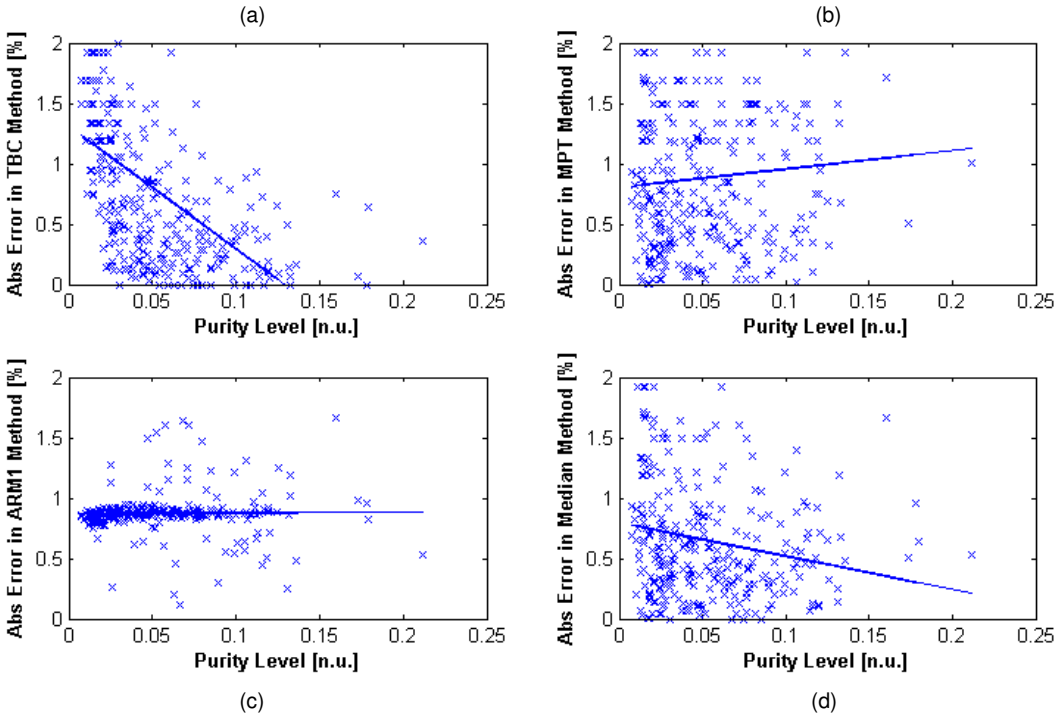


Figure 5-19. Purity level of subject lung sound recordings.
a) Purity level versus Abs % Error in TBC method, best fit line y=1.309-10.105x with R^2=0.263
b) Purity level versus Abs % Error in MPT method, best fit line y=0.806+1.521x with R^2=0.007
c) Purity level versus Abs % Error in ARM1 method, best fit line y=0.868+0.106x with R^2=4.817e-4
d) Purity level versus Abs % Error in MEDIAN method, best fit line y=0.798-2.773x with R^2=0.0361

From Figure 5-19 for the TBC and MEDIAN methods, higher purity recordings were a likely indicator for lower error rates. The linear fit for the TBC method had the highest correlation value (R^2 = 0.263) with a positive gradient, and from observing its plot in Figure 5-19a, a purity threshold of 0.7 could possibly be suggested. However, the MPT and ARM1 methods showed no significant change in error rates as purity increased because the data were widely scattered in Figures 5-19b and c. This suggests that purity level was not applicable to all signal processing methods for data

acquired with this mobile phone collection protocol. In fact, the noisy nature of the signal means that multiple frequencies were often present in the signal.

With heart rate estimation from heart sound data, the higher the BMI of an individual, the higher the estimation error. However, for lung sound recordings, this relationship was not observed as strongly (see Figure 5-20). The TBC, MPT, and MEDIAN methods hinted at slight increases in error levels for higher BMI subjects, but the spread of data points were very large and the correlation values were low. In Figure 5-20c, the accuracy of the ARM1 algorithm appeared to be independent of BMI.

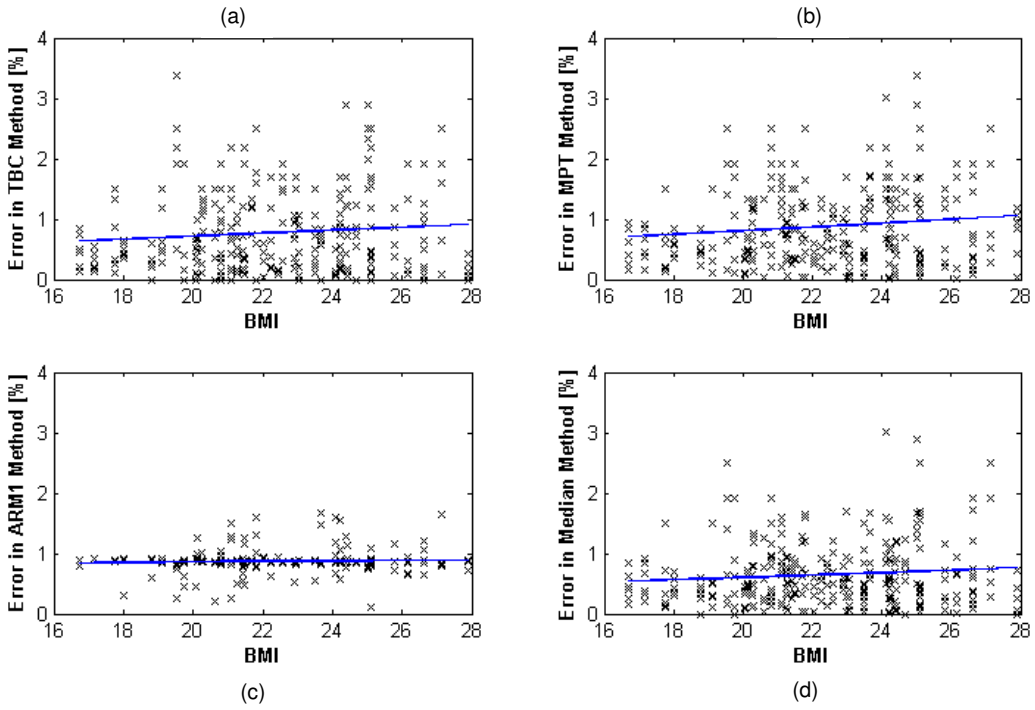


Figure 5-20. BMI vs. error rate in lung sound recordings.
a) BMI versus Abs % Error in TBC method, best fit line y=0.224+0.249x with R^2=0.009
b) BMI versus Abs % Error in MPT method, best fit line y=0.189+0.031x with R^2=0.017
c) BMI versus Abs % Error in ARM1 method, best fit line y=0.794+0.004x with R^2=0.003
d) BMI versus Abs % Error in MEDIAN method, best fit line y=0.225+0.019x with R^2=0.009

Lastly, the data was segmented according to healthy and unhealthy test subjects to study the influence of actual respiratory rate and purity level on estimation error. From Figure 5-21 for healthy subjects and Figure 5-22 for unhealthy subjects, the results from all the methods were plotted together for comparison. The best fit lines for the best method (MEDIAN) were plotted and shown in red. It was observed that regardless of the subject's health condition, higher respiration rates led to lower error rates, as seen in Figure 5-18. The correlation values for this trend were R^2 = 0.230 for healthy subjects and R^2=0.311 for unhealthy subjects. There was little distinction between the two groups of subjects, although it was noticed that there was slightly lower purity in signals for healthy subjects (see the cluster of data points at the lower end of the purity scale in the top graph of Figure 5-21).

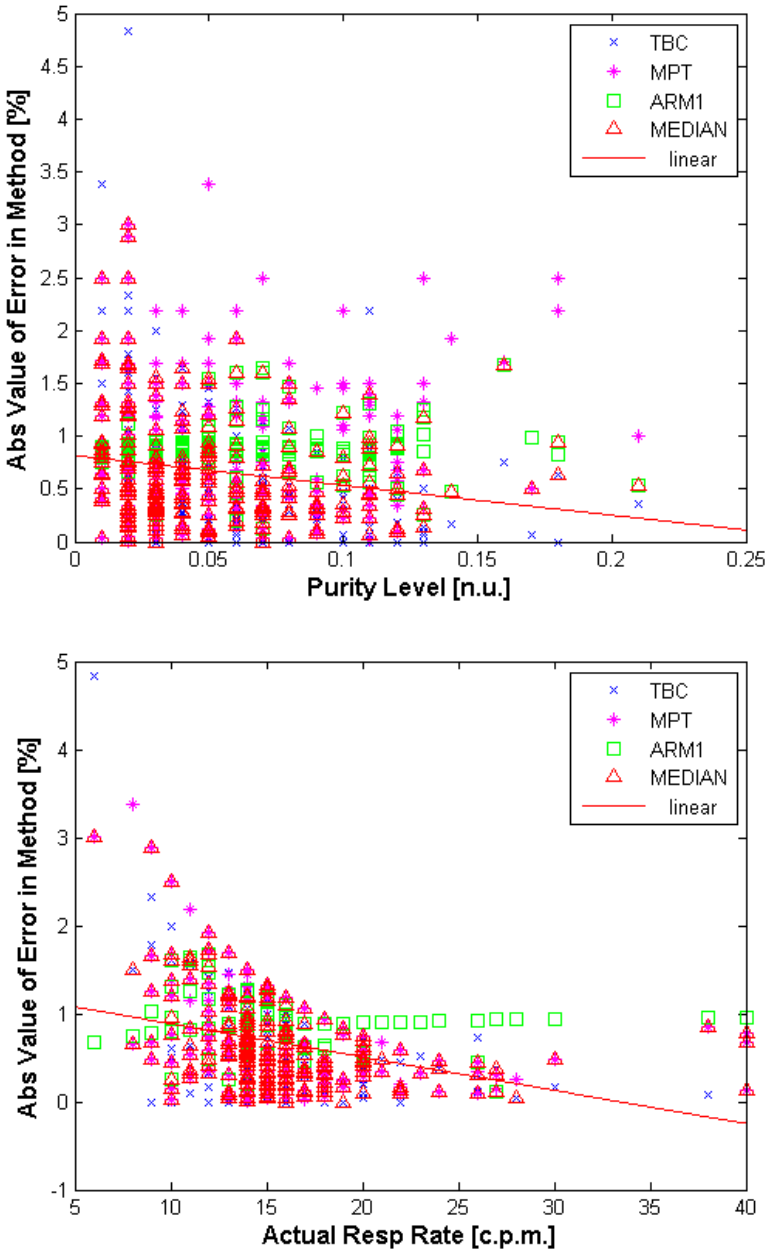


Figure 5-21. Healthy subjects and estimation errors from different methods. (Top) Best fit line for MEDIAN data is y=-2.800x+0.820 with R^2=0.015. (Bottom) Best fit line for MEDIAN data is y=-0.038x+1.300 with R^2=0.230

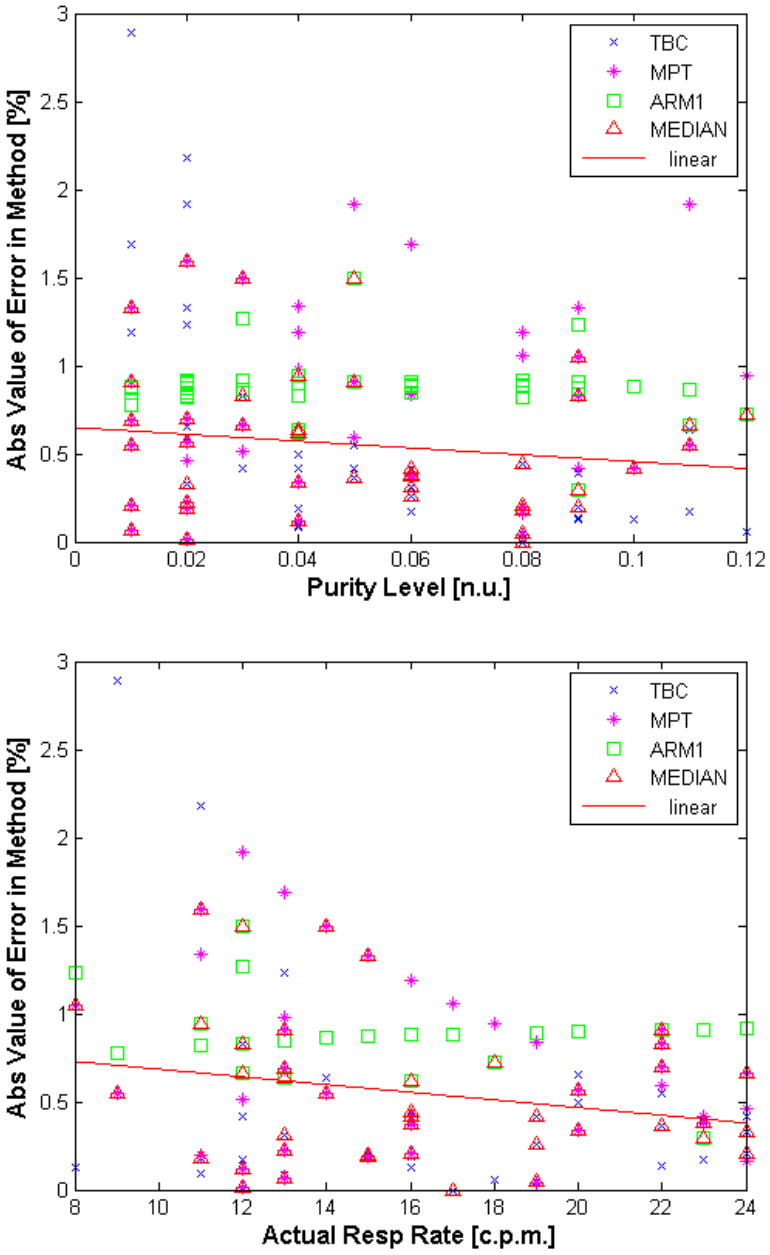


Figure 5-22. Unhealthy subjects and estimation errors from different methods.
(Top) Best fit line for MEDIAN data is y=-1.900*x +0.650 with R^2=0.041
(Bottom) Best fit line for MEDIAN data is y=-0.022x+0.900 with R^2=0.311

5.4 Discussion

The work in this chapter demonstrated the feasibility of recording heart and lung sounds using mobile phones with sufficient quality for analysis. Automated heart and respiratory rate estimation were successfully performed on the data with knowledge acquired about the limitations of current signal processing techniques. For example, the HR_{est2} method depended on dynamic adjustment of the peak triggering threshold, but it was unclear whether further changes in the threshold could have led to more optimal estimates.

One area of weakness was the need for a robust signal quality measure. It may be possible that a simultaneous feature extraction and classification approach based on wavelets would be promising (Martinez et al. 2000). Wavelets can provide a natural method for representing the short transients in heart and lung sounds and therefore may allow classifiers to distinguish between noisy, normal, and abnormal lung sound recordings. Recent work in this research area has been done by Hughes et al. (2003) and Schmidt et al. (2010), who used hidden Markov models for segmentation of heart sounds.

Another remaining challenge is to reduce noise, which will be an even larger problem in practice when CHWs use the device in busy clinics. Band pass filtering was applied to the data in this analysis to reduce high frequency noise. However, much of the noise remains in-band. Alternatives for dealing with in-band noise include adaptive filtering (which requires an exogenous input) or independent component analysis (ICA). The latter technique assumes a linear, stationary mixing of sources (noise and signal), which is probably true for the audio data being analyzed here. If the reflections are

relatively insignificant, ICA can prove to be a powerful technique. However, ICA usually requires at least as many sensors as independent sources. In other words, for two sources (heart or lung sounds and background noise), two microphones are required. Relatively few phones (except the Motorola Droid) are able to record in stereo. However, single-channel ICA algorithms do exist, and Jimenez-Gonzalez and James (2009) recently demonstrated (on a limited data set) that fetal heart sounds can be extracted from a single microphone. Although this method is promising, significant work needs to be performed to ensure that sensitivity to filtering parameters (such as the embedding dimension and delay time) are investigated in different noise conditions. Moreover, in non-ideal recording conditions where the signal-to-noise ratio is lower (as in the mobile environment), such methods may not be sufficient.

The study presented in this chapter represents a limited data set. A broader set of data should be collected in a larger study population which includes more unhealthy respiratory and cardiovascular patients. In terms of demographics, more females, elderly, and pediatric patients should be represented (the subjects in this work were primarily in their twenties). Despite these limitations, the results in this chapter appear to hold promise for low-cost automated lung sound analysis on a mobile phone.

Chapter 6

Mobile data collection framework for labeling signals and images

6.1 Introduction

Some of the limitations in development of the signal processing methods in this thesis were the noisy field environments and low number of heart and lung sound recordings available, particularly from unhealthy patients. In order to build a representative gold standard labeled database of these sounds, a robust and user-friendly infrastructure is required. Perhaps the cheapest and most effective platform would be one that leverages the portability and sophistication of mobile phones to collect and upload field data.

Within developing countries, mobile phones are attractive as a way to remedy the relatively little health informatics infrastructure that currently exists (Blaya 2008, Eng 2001). In an effort to take advantage of the low-cost and ubiquitous cellular network connectivity, more and more ad-hoc telemedicine solutions are being created for use in the field. EpiHandy (2008) is one mobile phone tool (based on J2ME) used for collecting survey data easily (i.e. on malaria or TB) by filling out forms and submitting them to a centralized server via SMS, TCP/IP, or Bluetooth. Nacer, a telemedicine solution by Voxiva in Peru, is aimed at improving maternal and child health in by having CHWs

submit patient information through the web or voice calls to receive medical expertise back. In Uganda and Malawi, FrontlineSMS:Medic (2009) employs two-way text messaging on mobile phones to allow texting to a large group of people (i.e. patients or CHWs), uploading of patient data to a medical record system, sending information for a remote diagnosis, or viewing a patient's electronic health record. In Mali, Pesinet employs a text-based system to track the weights of newborn babies (Menchi 2008).

The primary disadvantage of existing systems described above is that they were designed with very low bandwidths in mind, particularly in the case of SMS-based applications and are unable to cope with data loss over GPRS or WiFi. Unfortunately, such systems are impractical when one considers the nature of diagnostic audio data. As discussed in Chapter 4, the compression methods used in cellular telephony or soundcard characteristics of low-end phones are likely to introduce significant distortions to the diagnostically relevant frequency bands of the captured audio data. Therefore, standard software for capturing or transmitting audio data cannot be used, and an uncompressed audio capture algorithm on the mobile phone is required. The other disadvantage of the above systems is that they are proprietary, or make use of proprietary technologies or medical record systems. The lack of openness of such systems is a barrier to adoption and prevents effective auditing and longitudinal record-keeping (since the proprietary format will inevitably cease to be maintained and importing of the data into another database is likely to be prohibitively expensive).

Sana[39] is a robust platform with the potential to transmit bioacoustic data *accurately* to remote specialists for diagnosis. In Sana, patient data is transmitted via 3G, GRPS, or WiFi networks (instead of through cellular voice calls or SMS) to a back-

[39] http://sanamobile.org/

end medical record system for storage and easy retrieval at a later time. In this chapter, the functionality of Sana is extended to allow for the audio capture of heart and lung sounds. The additional features include: 1) a method to upload and review uncompressed audio data to the server, 2) an OpenMRS module to interact with rich media files, and 3) a method for uploading numerically coded medical ontologies to label patient data in a standardized format.

Deployment of a technology platform such as Sana in the field requires preparation aside from development of the technology itself. For example, Section 6.3.7 describes the system documentation that was written to aid others in the installation and IT maintenance of Sana. The documentation also includes a discussion of issues to consider concerning patient data security when implementing Sana to protect the patients' privacy, health and safety. The Sana platform was evaluated in terms of compliance with the Health Insurance Portability and Accountability Act (HIPAA) Security law (U.S. Department of Health and Human Services 2003). While HIPAA only applies to data recorded in the U.S., the barriers to adoption for Sana would be reduced if the system was shown to be compliant with HIPAA, which is considered stricter than the security and privacy laws in other countries. Section 6.3.8 then describes the experiences from a trip to the Philippines for preliminary investigative work on a pilot for Sana. Observations are reported concerning the obstacles that must be overcome in order to achieve successful adoption of a low-cost medical diagnostic technology in a resource-poor setting.

6.2 Workflow Overview

As introduced in Section 1.1, Sana is an existing end-to-end platform solution for remote medical diagnostics on a mobile phone, whereby a CHW can upload patient data to OpenMRS and receive a diagnosis from a specialist via SMS. One of the key advantages of Sana is the facility for a doctor to fully customize the workflow without the need for advanced software development skills. Although support for other platforms is planned, a CHW currently interacts with Sana through an Android client application, which contains a collection of procedures defined by doctors. Each procedure is intended to guide a CHW through a patient consultation to collect the appropriate clinical information in order to provide accurate and timely data for the remote doctor.

Sana was designed with a generic data transmission facility in mind. As such it can upload data of any format to the back-end medical record and alert users that the data are available for review. Therefore, Sana is an ideal system to provide secure storage and annotation of heart and lung sounds captured from a mobile phone. Moreover, the Sana platform can naturally be employed to provide medical expertise to patients who present with lung or cardiovascular symptoms (such as cough, shortness of breath, or tachycardia). During a consultation of a patient with a CHW, the patient is first authenticated via a barcode (scanned using the phone's built-in capabilities) or through the patient name and demographics. A remote query (or local query of a small cached database) provides the CHW with the ability to decide if this patient already exists in the medical record system or if a new patient should be created. Using a custom procedure on the phone involving a series of questions and prompts, the CHW collects information on patient history, physically examines the patient, and can then

capture audio recordings of the patient's heart or lung sounds. After the customized form has been completed, the data is automatically (asynchronously) uploaded to the medical record for the patient. If no connection to the remote database is available, then the Android application stores the data for later automatic upload when a connection becomes available.

Once all the data related to a given CHW-patient encounter (or visit) is uploaded, the remote specialist is automatically notified of new data to be reviewed. The specialist logs into the OpenMRS standard web interface (accessed through any web-capable device such as a laptop or mobile phone) to review the patient's symptoms and then listens to the heart or lung sound recordings. The specialist then assigns a set of diagnoses to the patient from a prescribed (pre-loaded) medical ontology, such as SNOMED CT[40] (Systematized Nomenclature of Medicine Clinical Terms) (UMLS 2010). Then any treatment recommendations, comments, or an urgency level are added to the case. At this point, the data can either be automatically sent via SMS (and email) back to the referring CHW, or the server-side application can wait for additional experts to provide their annotations of the data so that a more accurate consensus decision can be made.

There is also the possibility that the specialist will request more information concerning the patient such as lab tests or a chest X-ray if the differential diagnoses include pneumonia. If an X-ray is recommended, then a digital image of the X-ray can be captured with the phone's camera and uploaded to the patient's centralized medical record using Sana. Again, the expert will review the relevant patient history and media files, make a definitive diagnosis, prescribe treatment, or refer the patient to a nearby

[40] www.ihtsdo.org/

hospital for more medical attention. With this workflow, it is evident that Sana is able to integrate a patient's clinical information with heart and lung sound recordings, as well as images such as X-rays, in order to build a complete medical record for patients within OpenMRS.

6.3 System Description

The Sana platform described in Sections 6.3.1 to 6.3.3 already existed prior to this thesis and was created by the team of Sana developers (except for the *Shortness of Breath* procedure used as an example in 6.3.1). The work in this thesis contributes the features described in Sections 6.3.4 to 6.3.8 to the Sana platform.

6.3.1 Mobile Phone Application and Interface

The Android phone application allows a CHW to collect appropriate data on patient symptom information based on a step-by-step questionnaire. Developed through consultation with experienced clinicians and relevant literature (Bates 1995), the *Shortness of Breath* procedure is shown in Figure 6-1 as a series of screenshots from the phone's interface. This procedure closely follows the protocol for collecting lung sound data in Chapter 5, and involves recording lung sounds across the patient's back at each of six locations: the upper left, upper right, center left, center right, lower left, and lower right back. In Figure 6-1, from left to right, the screenshots display:

1. The initial main menu, which allows the user to choose the appropriate interaction.
2. A list of available procedures including the *Shortness of Breath* evaluation.

3. A request for the relevant clinical history of the patient.
4. Instructions for the CHW to communicate to the patient on how to breathe deeply for the acoustic lung sound recordings.
5. Instructions on where to put the stethoscope device (from Chapter 4) for recording lung sounds from the initial location with an accompanying figure.
6. Instructions on the next location to record lung sounds.
7. The final screen allowing the user to begin the upload or return to previous pages to correct any information.

Note the progress bar across the top of the application indicating the percentage of the procedure completed so far. Note also that the data upload occurs in the background, so that the CHW can continue to evaluate other patients in the meantime.

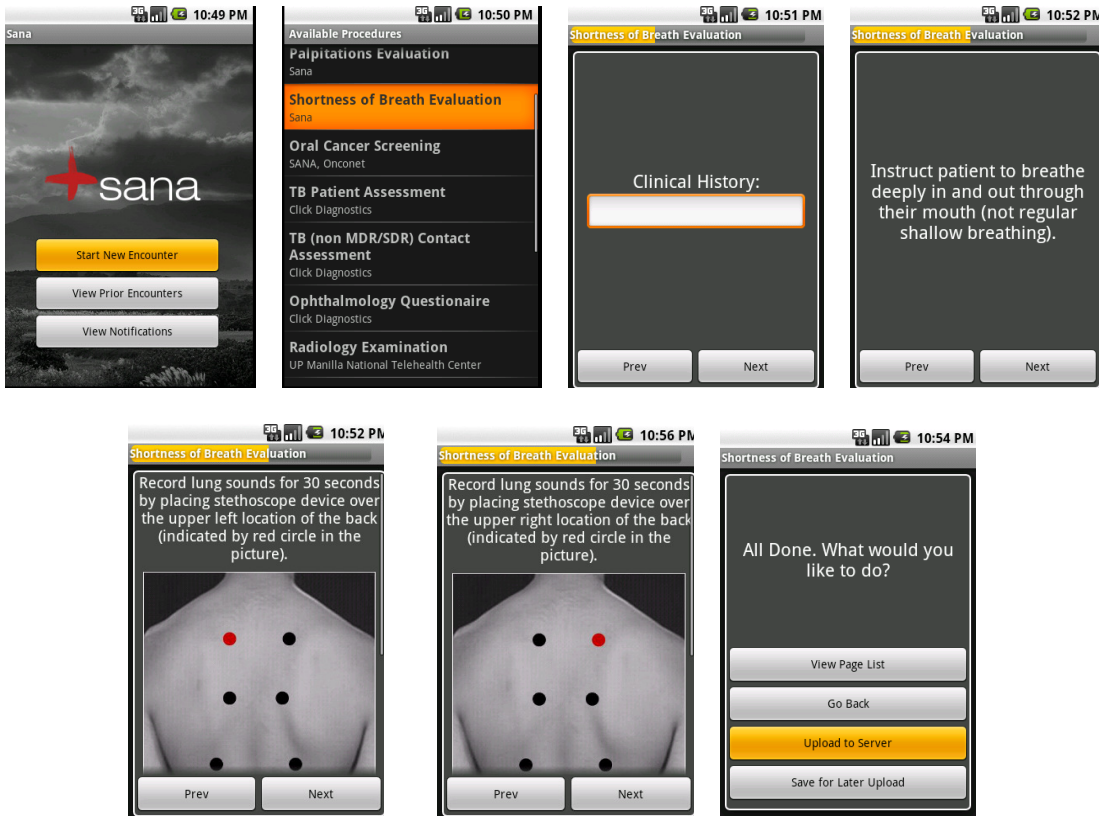


Figure 6-1. Screenshots of the *Shortness of Breath Evaluation* on the Sana Android application.

6.3.2 The Intermediate Mobile Dispatch Server

Once a patient visit is completed, the healthcare worker uploads the case to the intermediate Mobile Dispatch Server (MDS) to be subsequently sent to OpenMRS (2010). The MDS is a Django web server implemented in Python that serves as an abstraction layer between the phone handsets and the backend medical record systems. The phone communicates with the MDS using a JavaScript Object Notation interface[41] through HTTP requests over the encrypted cellular network. Future integration with operating systems other than Android (for example, with the Symbian operating system) should be relatively easy and EMR independent because the application on the new handset would simply use the application protocol interface (API) provided by the MDS, instead of integrating directly with the medical record system (such as OpenMRS). Similarly, the MDS could be extended to send patient cases to any other medical record system.

6.3.3 The Back-end Medical Record System: OpenMRS

The patient case is finally uploaded from the MDS to OpenMRS (2010) to provide specialists with an interface for reviewing information about Sana patients. OpenMRS was selected as the platform of choice because it is the leading open-source electronic medical record system used in developing countries and currently piloted in approximately 14 countries across the globe including Kenya, Peru, Haiti, and Malawi. Developed by collaborators from Partners in Health, the Regenstrief Institute, and the South African Medical Research Council, OpenMRS is an extensible and scalable

[41] JSON - http://www.json.org/

platform based on a unique data model that incorporates concepts, clinical observations, patients, providers, encounters, forms, and drug orders (OpenMRS 2007).

The Sana module were developed as a plug-in to the OpenMRS platform in order to integrate with the MDS and display a queue of pending patient cases requiring review by a medical specialist.

6.3.4 Rich Media Content Viewer in OpenMRS

In order for heart and lung sound recordings to be played and supplemented by relevant clinical information, the capacity to review media files and patient cases within OpenMRS is required. OpenMRS previously did not have an easy-to-use interface for viewing rich media consisting of images, audio, and videos. These media files were stored as observations under a specific patient encounter and were accessed using a default observation viewer with limited functionality (see Figure 6-2). Images could not be manipulated for a more detailed and accurate view of the data, while audio or video files could not be previewed at all without downloaded them and using locally installed software. A former OpenMRS student intern attempted to build an interactive image viewer using the Google Maps API, but significant limitations still existed with this module because it did not support playback of audio or video files (OpenMRS 2009a). Furthermore, if the module was removed from the server instance, then the patient data would no longer be retrievable.

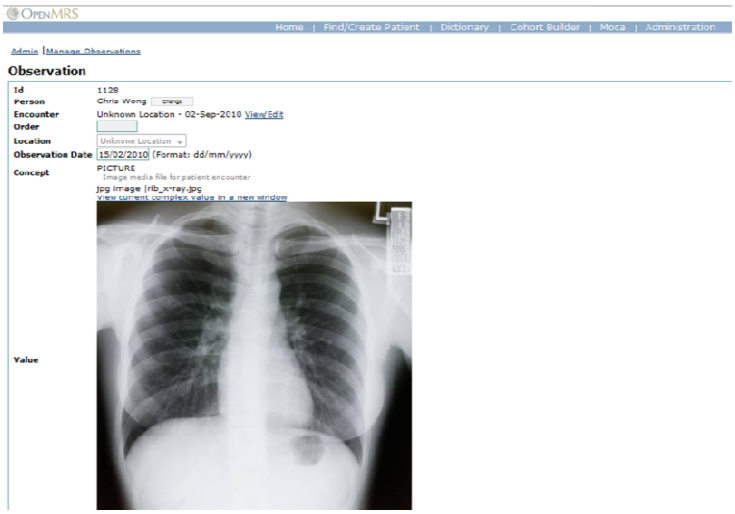


Figure 6-2. OpenMRS built-in observation viewer.

The built-in encounter viewer provided by OpenMRS also had potential for improvement (as seen in Figure 6-3). For a particular patient encounter, the viewer presented a list of text-based observations with two columns: the concept names (i.e. pain level, history, physical findings) and their corresponding values (i.e. value 10 for pain level, value one for number of previous miscarriages). This list of observations (or <concept name, value> pairs) did not provide a comprehensive view of the encounter in terms of previews for the image, audio, and video files. Each observation had to be clicked in order to download the media file or to learn more details.

Based on evaluation of the available functionality in OpenMRS, there was a clear need to develop an advanced media viewer, which would be fully compatible with the OpenMRS data model and offer an easy-to-use interface for specialists to review patients' symptomatic information (text) and media files (images, audio, video).

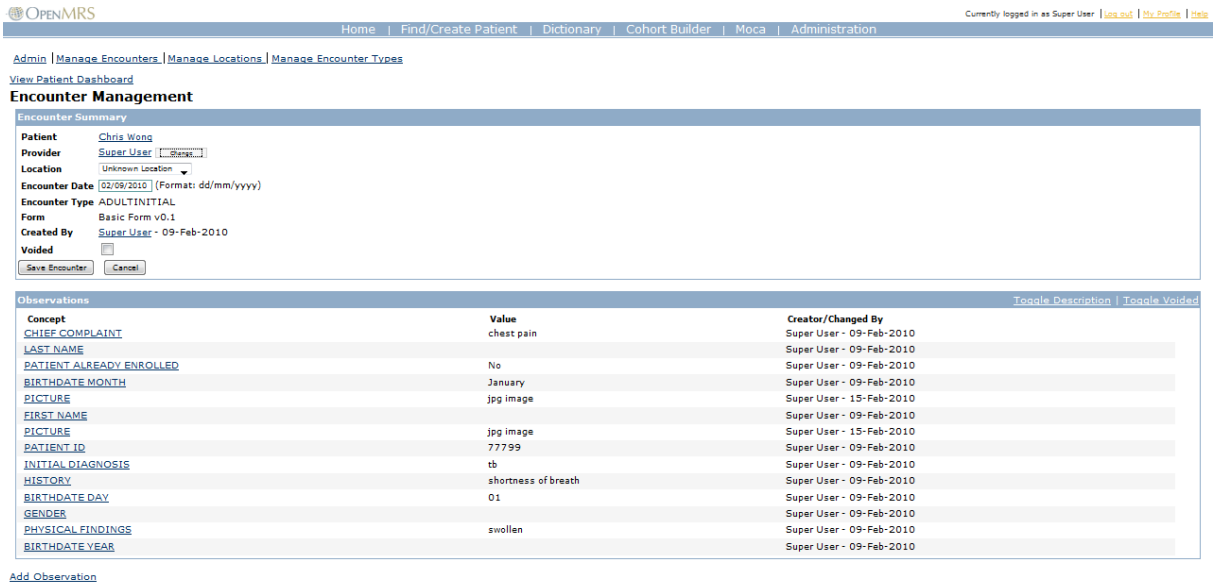


Figure 6-3. OpenMRS built-in encounter viewer with a list of all observations associated with one patient-provider encounter.

The media viewer was developed as part of OpenMRS Google Summer of Code 2009 under the supervision of Dr. Gari Clifford. It was designed as a stand-alone module from the existing Sana module to be contributed to the open source OpenMRS community. Then other users could add the media viewer module to their instances of OpenMRS for media support, without requiring them to use Sana as well (OpenMRS 2009b). This is in line with the design considerations to make the client and server independent entities.

A significant portion of the media viewer was built in Flex 3 because of the rich media experience afforded by the Adobe Flash Player.[42] The Flash media viewer plug-in could then be embedded into the OpenMRS (2009b) Media Viewer Module, other medical record systems, or be stand-alone on the web. The class diagram for the implementation of the Flash component is depicted in Figure 6-4. The Flash component

[42] http://www.adobe.com/products/flashplayer/

communicated with the OpenMRS server through Java servlets to retrieve information on specific encounters and observations within those encounters.

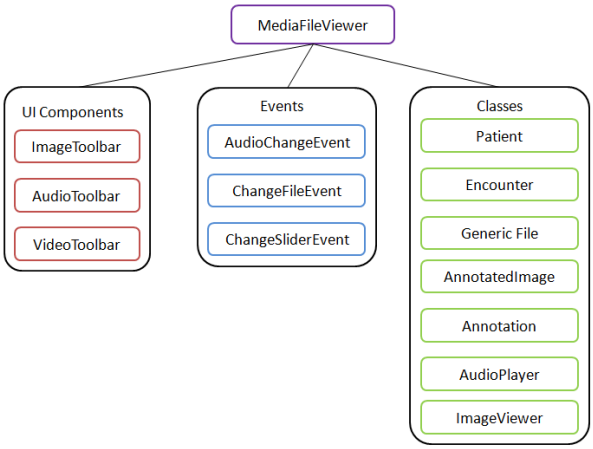


Figure 6-4. Class Diagram of Media Viewer Flex Component.

The media viewer allows users reviewing patient case to zoom and pan within the image, in addition to manipulating its brightness (Figure 6-5), contrast (Figure 6-6), and sharpness (Figure 6-7) with slider controls. If a physician needs to undo any changes to the image, there is also a convenient button to revert to the original image. Full-screen viewing and the option to download the file are available too.

In addition to displaying images, the media viewer can also play MP3 audio files and FLV videos (Figure 6-8). The controls for each type of media (Image, Audio, or Video Toolbars) are automatically displayed on the bottom of the screen according to the type of media file selected in the row of file thumbnails across the top of the viewer. Support for uncompressed audio playback (as required for lung sound diagnostics) has not yet been implemented because it is currently unknown if a library can stream audio without distorting compressive techniques. Heart and lung sounds review therefore requires the user to download the WAV file locally for playback.

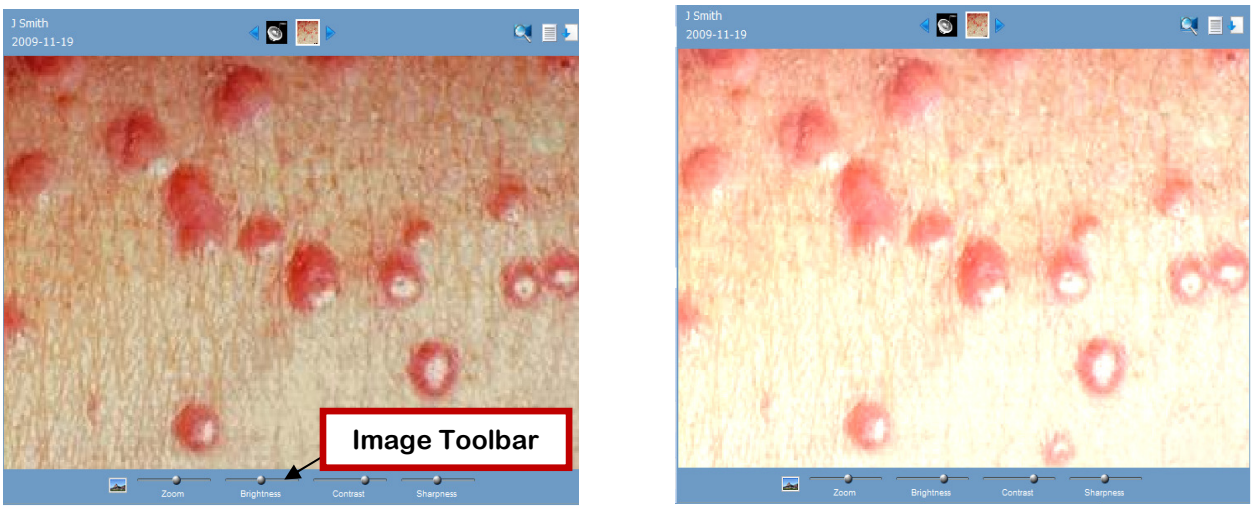


Figure 6-5. Change brightness of an image in the media viewer.

Figure 6-6. Change contrast of an image in the media viewer.

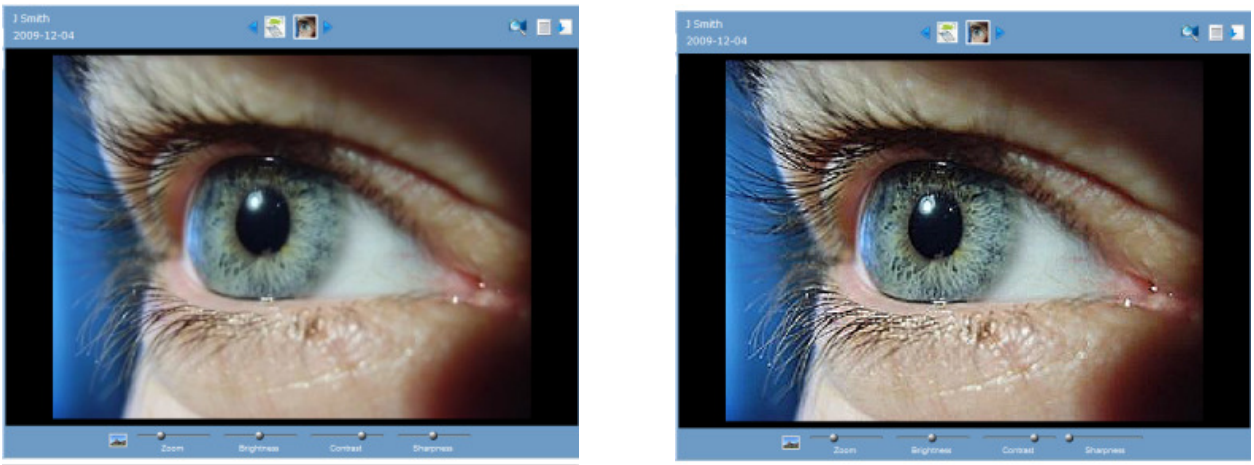
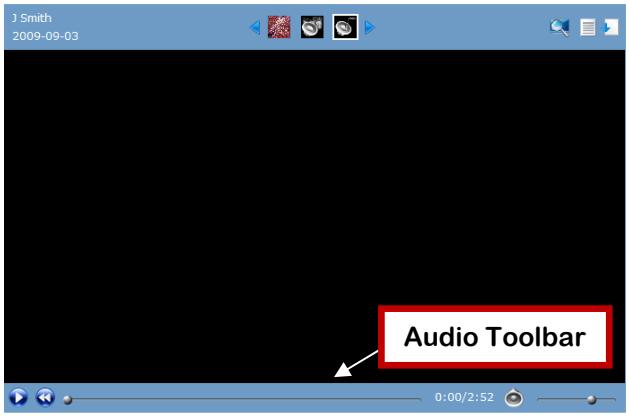


Figure 6-7. Change sharpness of an image in the media viewer. (Focus on eyelashes)



(a)

(b)

Figure 6-8. Playing a) an MP3 audio file and b) FLV video file in the Flash media viewer.

In OpenMRS, the Sana module was integrated with the Flash Media Viewer module for a comprehensive interface to review patient cases seen in Figure 6-9. The Media Viewer Flex component (embedded within the Flash Media Viewer Module of OpenMRS) is shown on the right half of Figure 6-9. A custom panel (on the left half of Figure 6-9) was built into the Sana module and implemented using HTML, Javascript, and CSS. Figure 6-9a shows basic demographic information on the patient and a link to the OpenMRS built-in patient dashboard to view past encounters. Part b of the panel in Figure 6-9 displays the question and response pairs (for the procedure on the Sana phone application) to detail the patient's clinical history and symptomatic condition. The last section (Figure 6-9c) prompts the specialist for an impression of the patient case

with fields to input patient diagnoses, treatment recommendations, comments, and urgency level.

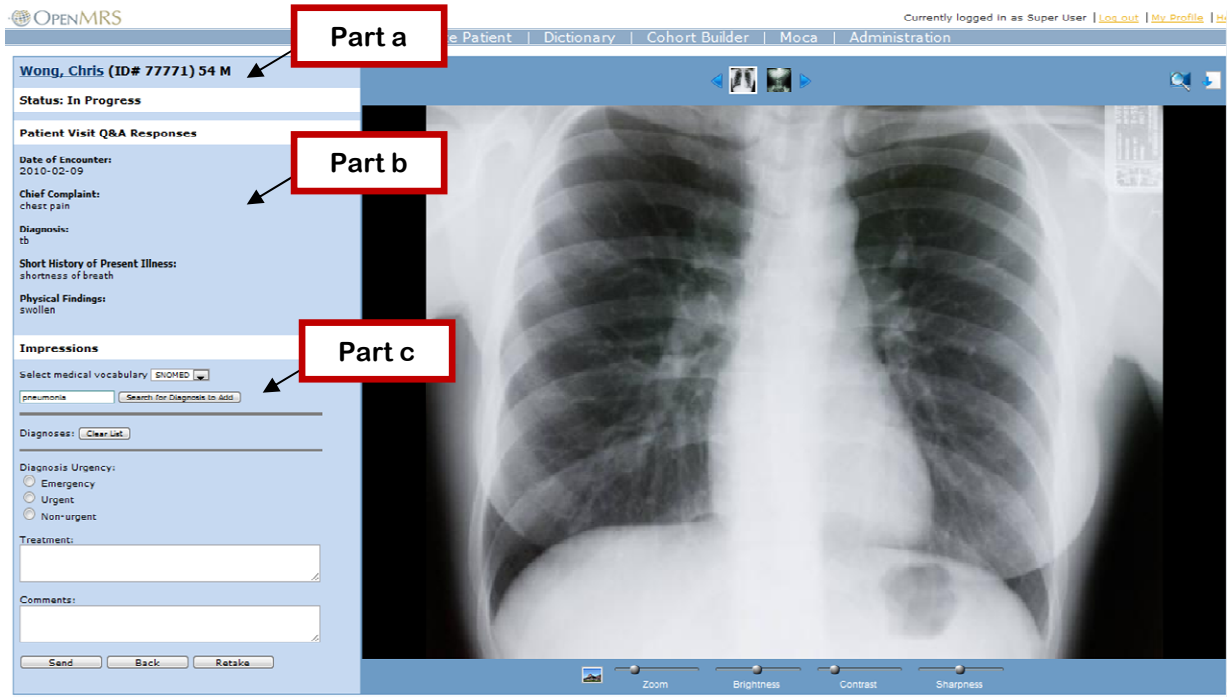


Figure 6-9. Screenshot of the working version of the media viewer module in OpenMRS.

After the information is filled out, the specialist sends the diagnosis and recommended plan back to the referring clinician, after which the case is marked as closed. For any closed case, one can always return to the case to review what diagnoses was prescribed in the past (see Figure 6-10).

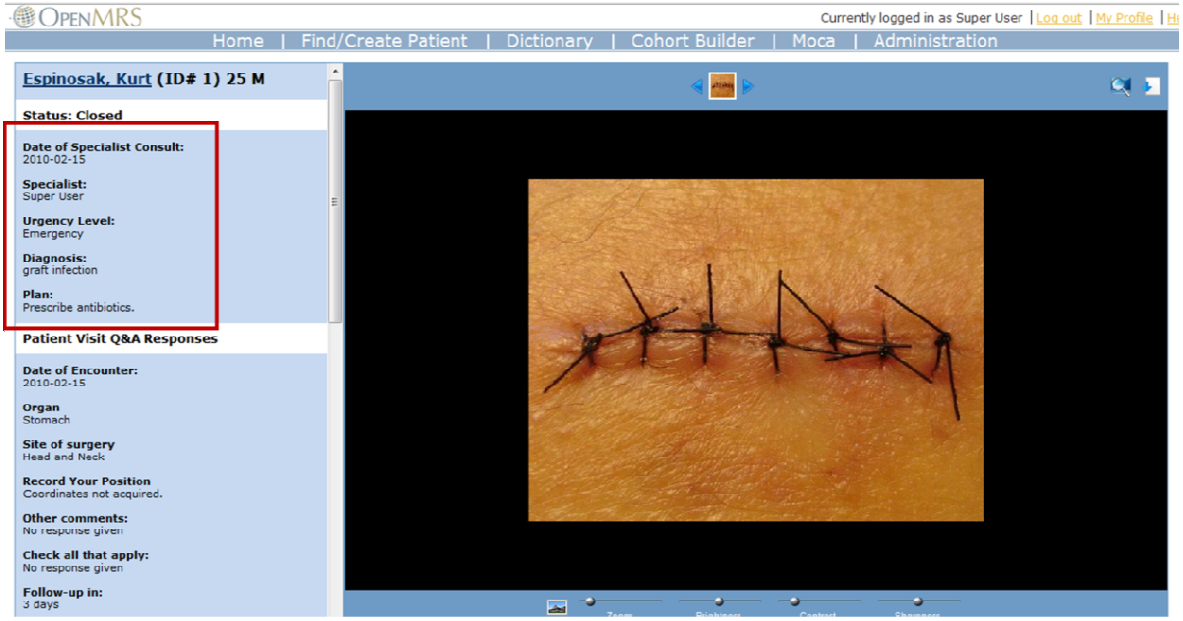


Figure 6-10. Diagnosis made (within red rectangular box) and related information for a closed patient case. (Red rectangular box not present on actual interface).

If the originally-provided information from the phone upload is insufficient to make a reliable diagnosis, there is a button at the bottom of the panel to request additional information. The specialist can request a retake of the image, additional media files, or more details about the patient. The specialist is then prompted with a free text box to explain the specific request to the referring clinician, which is delivered via email and SMS. Then the patient case is marked as “deferred” until more information is acquired. Specialists can access a list of all the pending, deferred, or closed cases in OpenMRS using a filtering option available on the Sana queue webpage.

6.3.5 Medical Ontologies and Labeling Data

A standardized approach is required for the labeling of heart and lung sound audio files with appropriate findings and associated diagnoses. Since free text can easily be mistyped, involve abbreviations or variable structures (see Figure 6-11),

specialists are asked to assign patients diagnoses using a standardized medical lexicon. Sana was customized to allow for integration with any medical lexicon, so that doctors had to choose from a predetermined vocabulary of diagnoses. Examples of these lexicons include but are not limited to SNOMED CT, LOINC[43] (Logical Observation Identifiers Names and Codes), and ICD-10[44] (International Classification of Diseases). These vocabularies assign numerical codes to diagnoses, findings, procedures, and pharmaceutical drugs in hierarchical ontology trees (UMLS 2010).

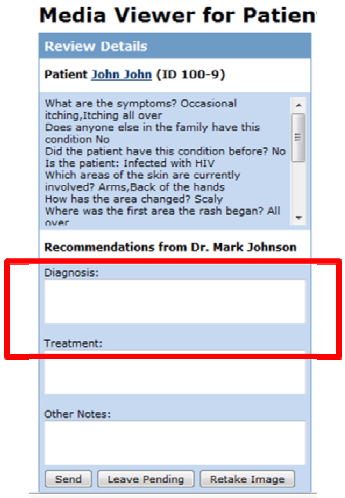


Figure 6-11. Free-text entry for diagnoses in initial design of the Sana encounter viewer.

Users can still add supplementary information through free text, but this is just for later human review, or possible future natural language processing (NLP). However, NLP of medical text is a difficult problem and can lead to serious errors. Moreover, it is extremely difficult to compare different annotators to reach a consensus diagnosis if a limited set of standardized alpha-numeric codes are not used.

The functionality of allowing users to upload medical lexicons into OpenMRS was a feature developed in collaboration with Drs. Cleo Maehara and Richard Loomis as

[43] http://loinc.org/
[44] http://www.who.int/classifications/icd/en/

part of the Fall 2009 MIT 6.872 Biomedical Computing course (MIT OpenCourseWare 2003b). The lexicons were integrated with the OpenMRS built-in concept dictionary to maintain consistency with the OpenMRS data model. OpenMRS (2008) uses concepts to encode and categorize observations on patients, so concepts can be medical tests, diagnoses, findings, or drugs. For example, a patient observation requires a concept such as “blood pressure” and value such as “120/80.” Each concept also has a concept class, so the pneumonia concept would have concept class “diagnosis” and could be added to the medical record of a patient who had pneumonia. As seen in Figure 6-12, the concept “pneumonia” (provided by the default OpenMRS concept dictionary) did not correspond to a numerically coded ontology, which would make data mining and comparisons difficult. Fortunately, the concept dictionary is extensible so that users can add new concepts of interest. The web form for manually adding individual new concepts to OpenMRS is displayed in Figure 6-13. This method would be inefficient however for a medical vocabulary such as SNOMED CT with hundreds of thousands of terms (UMLS 2010).

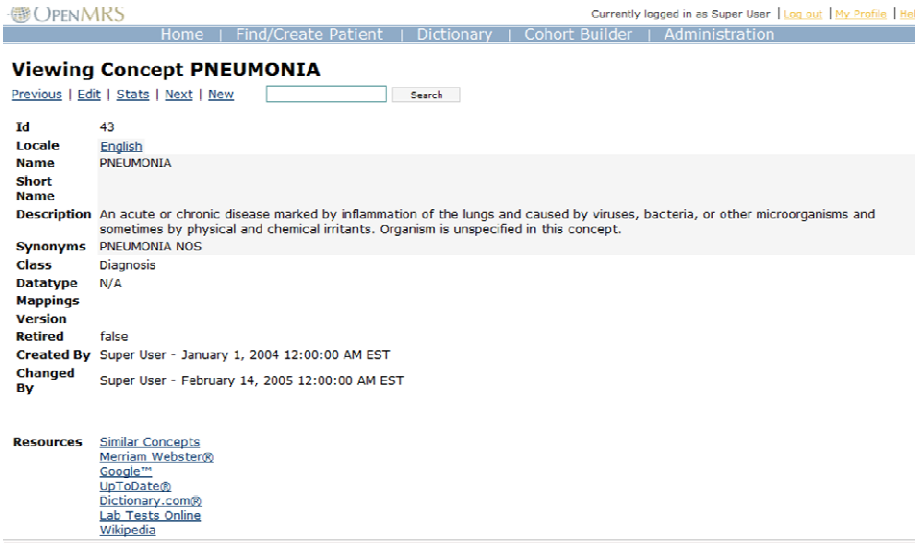


Figure 6-12. Pneumonia concept (without numerical code) that comes with the built-in OpenMRS concept dictionary.

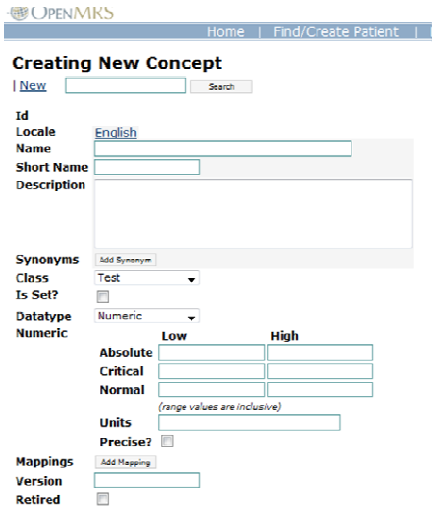


Figure 6-13. Form to create one new concept in OpenMRS.

A new graphical interface was implemented as part of the Sana OpenMRS module in order to allow new concepts to be efficiently and flexibly added *en masse* (See Figure 6-14). The OpenMRS instance was updated so that the administrative user could add a new medical lexicon (called a concept source in OpenMRS) by providing the concept source with a new name, description, and CSV (comma separated values) ASCIII file. The file contained a list of 1) the concept names for the terms in plain English, 2) the alpha-numeric ID number, and 3) concept class. See Tables 6-1 and 6-2 for examples of the three-column CSV file. During the concept creation process in OpenMRS, real-time updates are displayed to the user on the progress of the lexicon import process. In addition, helpful error messages (i.e. reminders of which fields are required in the upload form) are provided to the user with instructions on potential system settings that must be changed before the process can be completed successfully.

The administrative user can also update the medical lexicon at any point in time. When later versions of a medical vocabulary are released, then the user only needs to

obtain separate lists of the retired and newly added terms[45]. Usually the organization issuing the vocabulary (i.e. International Health Terminology Standards Develop Organization for SNOMED) will publish a list of changes, so a user could manipulate a spreadsheet file to obtain these two lists. In the case where the user wants to completely remove a medical vocabulary, the interface also has an option to delete existing concept sources, which will void the concept source in the OpenMRS database so that all the concepts belonging to that concept source are no longer valid or searchable.

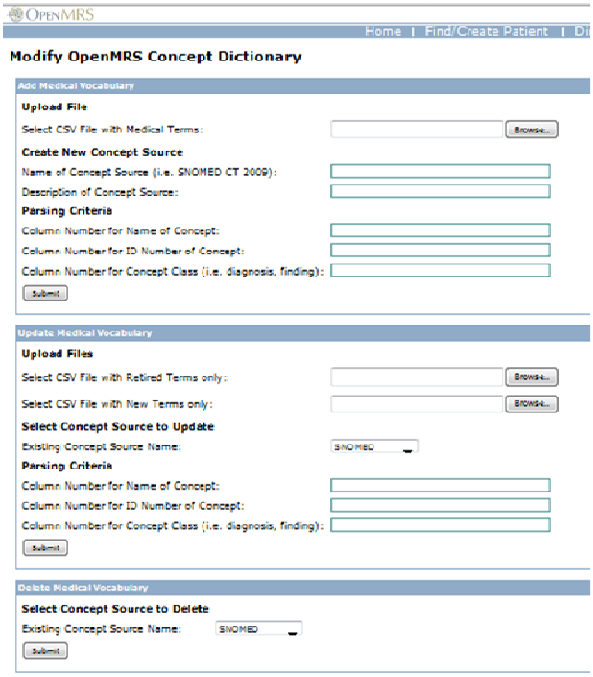


Figure 6-14. New user interface to manage medical vocabularies directly integrated with OpenMRS Concept Dictionary.

For demo purposes, the medical vocabulary loaded into the Sana demo server was the Veterans Health Administration / Kaiser Permanente Problem List Subset of SNOMED (UMLS 2008), which was added to OpenMRS under the “SNOMED” concept source. This lexicon contains 16,000 terms focused on common diagnoses and

[45] Although clearly changing the entire lexicon requires that old concepts be mapped to new concepts.

findings. An example SNOMED concept added into OpenMRS using the Sana module is in Figure 6-15, where the SNOMED ID is noted in the red box. The concept dictionary is searchable and the newly added SNOMED concepts also appear in the search results (see Figure 6-16).

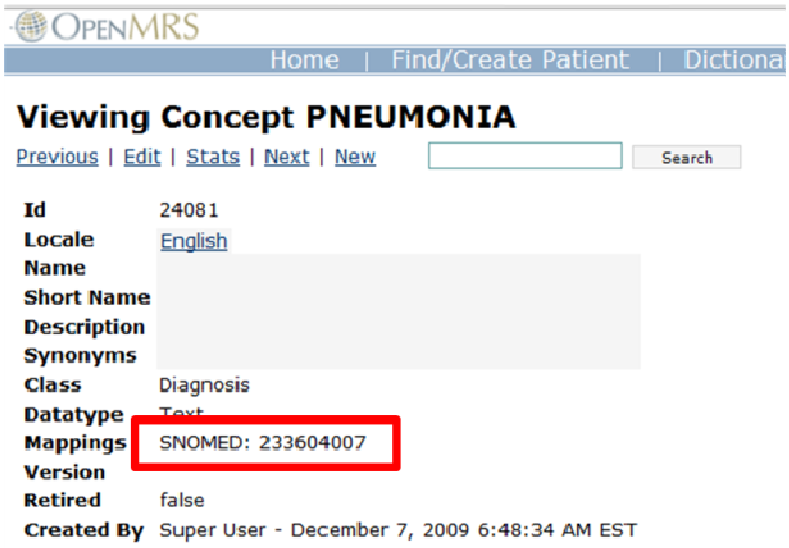


Figure 6-15. Concept pneumonia with concept source SNOMED and SNOMED ID 233604007 as the mapping code.

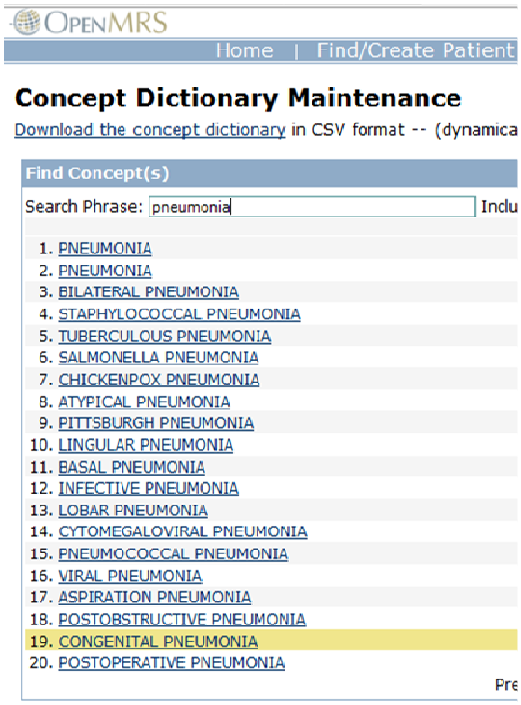


Figure 6-16. Searching the concept dictionary for “pneumonia” presents much more choices for diagnoses after the SNOMED concepts have been loaded into OpenMRS.

Since doctors are likely to assign different sets of diagnoses according to their specialty, the range of possible vocabulary terms for the diagnosis of a patient case can be narrowed. For example, radiologists may only need to assign diagnosis terms from RadLex,[46] which is a specific radiology lexicon. In the *Impression* section of the Sana encounter viewer, there is an option to limit concept searches to a certain lexicon (i.e. RadLex, SNOMED, or ICD-10) that that has already been loaded into the OpenMRS instance (see Figure 6-17). The "Default" option means that all concepts in the concept dictionary, regardless of lexicon, will be searched.

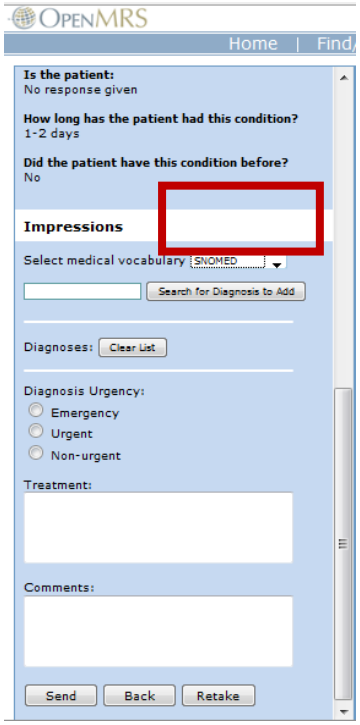


Figure 6-17. Option for which medical vocabulary to search for assigning diagnoses to a patient case.

After the specialist inputs a search string, the results appear as a list of links displaying the various coded concept names that match. The specialist can select the most appropriate diagnosis term listed in the search results to be added to a list of

[46] http://www.radlex.org/viewer

diagnoses tagged to the patient encounter. For a single patient encounter, multiple diagnoses can be made by the same or different specialists (see Figure 6-18). After the entire impression is submitted to the server, a new observation is added to the encounter for each diagnosis made by the specialist. See Figure 6-19 for the diagnoses (and their SNOMED ID numbers) added to the patient's encounter.

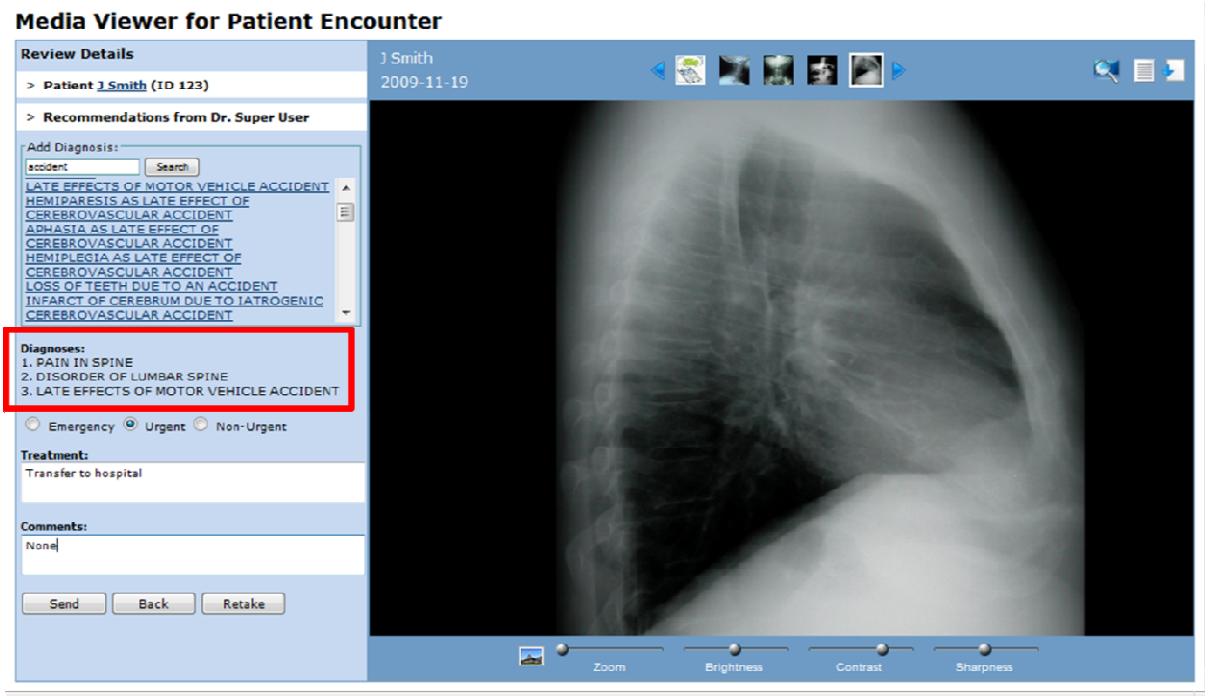


Figure 6-18. Assigning multiple diagnoses to a patient case in the Sana encounter viewer.

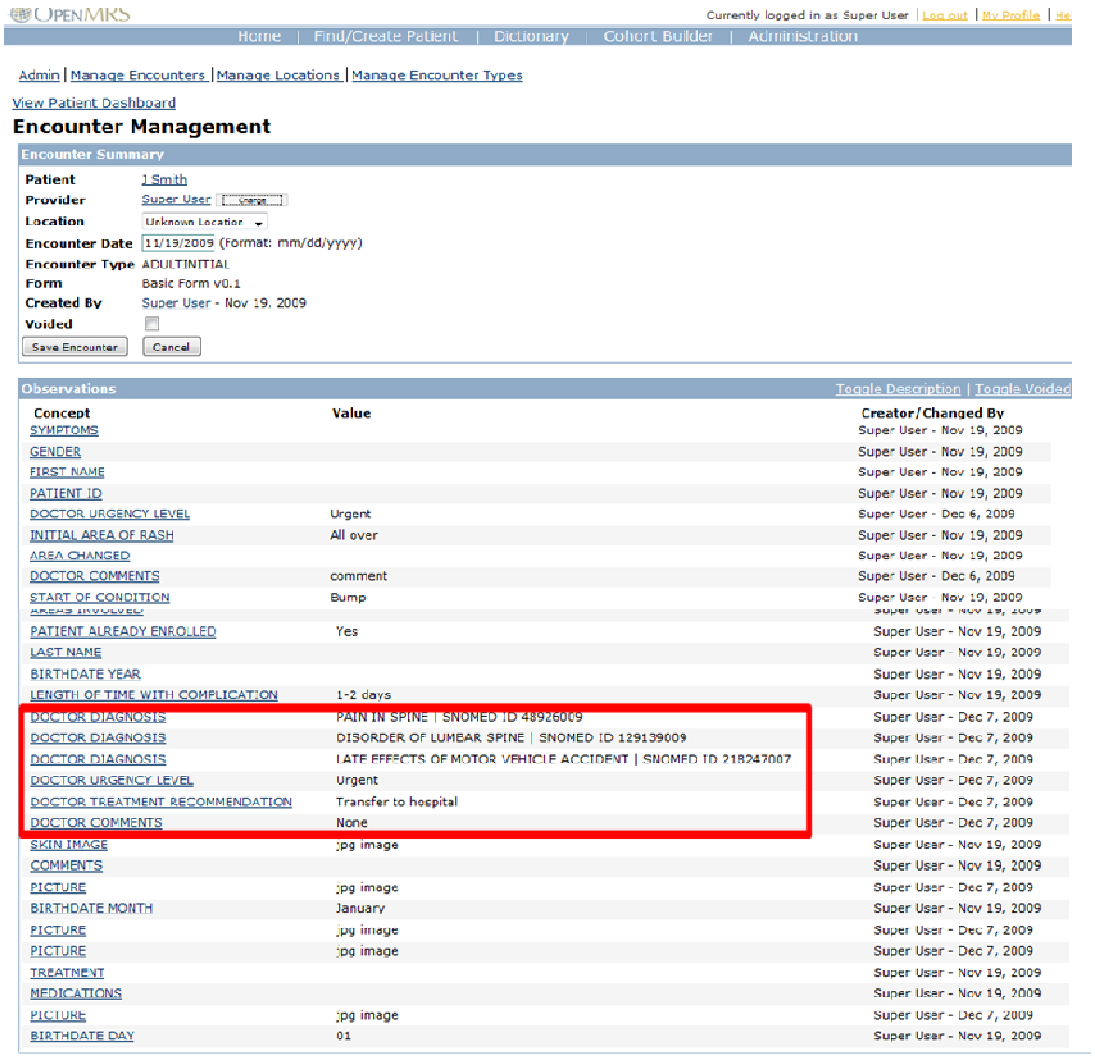


Figure 6-19. Updated patient encounter after specialist has reviewed case and tagged it with SNOMED diagnosis codes.

6.3.6 Audio Upload and Review in Sana

Two new workflows on heart and lung auscultation, respectively, were added to the Sana phone application's repository of procedures. The assumption was that the CHW would have access to a stethoscope device integrated with the mobile phone (similar to the one described in Chapter 4). As seen in Figure 6-1, the *Shortness of Breath* evaluation involves collecting patient symptoms, clinical history, and lung sound recordings from six locations on the back of the patient. The *Palpitations* evaluation

similarly collects patient symptoms, relevant medical history, and two heart sound recordings (from the mitral and pulmonic locations on the chest of the patient).

In order to upload uncompressed audio WAV files to OpenMRS, the Sana features to upload binary files from an Android client (built by Zack Andersen) and to receive binary files at the MDS (built by RJ Ryan) were slightly modified. In the resulting prototype, a CHW would go through the auscultation procedures described above. In order to record uncompressed WAV files, the CHW would have to launch the Hertz WAV Recorder[47] Android application (available for free from the Android Market), record the subject's body sounds using a stethoscope attachment, and then return to the Sana application to upload the appropriate file from the SD card.

Once uploaded to OpenMRS, the heart and lung sound audio files can be downloaded from the Sana queue and played by the specialist on a local computer. The files must be downloaded because the browser's built-in functionality to playback audio may involve unknown compression and streaming methods that may distort the clinical content of the signals. Furthermore, the audio cannot be played in the Flash media viewer since Flash applications only support the MP3 audio codec but not the uncompressed WAV format.

After the specialist downloads and listens to the sounds with standard locally installed software and a set of headphones of reasonable[48] quality, she can use the Sana encounter viewer assign a numerically coded diagnosis to the patient's case. A medical lexicon of possible diagnoses and findings that can be detected during traditional auscultation is used to respond to the *Shortness of Breath* and *Palpitations*

[47] http://www.appstorehq.com/hertz-thewavrecorder-android-162806/app
[48] With a frequency response from 20 Hz to 10 kHz and low total harmonic distortion

evaluation cases. An experienced doctor from the Sana team outlined the possible terms that could be used to describe heart and lung sound data. The corresponding SNOMED CT (UMLS 2010) ID numbers for these terms were found and uploaded into OpenMRS using the interface in Figure 6-14 as the concept sources “Auscultation Findings” (see Table 6-1) and “Auscultation Diagnoses” (see Table 6-2).

SNOMED ID	Concept Name	Concept Class	SNOMED ID	Concept Name	Concept Class
248654002	Third sound gallop	Finding	47525004	Coarse respiratory crackles	Finding
4592006	Fourth sound gallop	Finding	62640008	Fine respiratory crackles	Finding
7036007	Pericardial friction rub	Finding	24612001	Wheeze - rhonchi	Finding
88610006	Heart murmur	Finding	248623001	Respiratory squawk	Finding
251151005	Arrhythmia	Finding	248624007	Respiratory squeak	Finding
			70407001	Stridor	Finding
			301274008	Inspiratory wheeze	Finding
			301275009	Expiratory wheeze	Finding
			301703002	Expiratory polyphonic wheeze	Finding

Table 6-1. SNOMED findings for heart (left half) and lung sounds (right half). Terms from UMLS (2010).

SNOMED ID	Concept Name	Concept Class	SNOMED ID	Concept Name	Concept Class
417996009	Systolic heart failure	Diagnosis	155582008	Mild chronic obstructive pulmonary disease	Diagnosis
418304008	Diastolic heart failure	Diagnosis	155583003	Moderate chronic obstructive pulmonary	Diagnosis
390722003	Aortic stenosis	Diagnosis	155584009	Severe chronic obstructive pulmonary disease	Diagnosis
155349009	Non-rheumatic aortic sclerosis	Diagnosis	390798007	Asthma finding	Diagnosis
194734000	Mitral stenosis and aortic insufficiency	Diagnosis	278516003	Lobar pneumonia	Diagnosis
194746001	Rheumatic pulmonary stenosis and	Diagnosis	64667001	Interstitial pneumonia	Diagnosis
194997002	Pulmonary stenosis, non-rheumatic	Diagnosis	155606001	Pleural effusion	Diagnosis
156915002	Atrial septal defect	Diagnosis	36118008	Pneumothorax	Diagnosis
30288003	Ventricular septal defect	Diagnosis			
251009000	Tricuspid regurgitation	Diagnosis			
194742004	Tricuspid stenosis, cause unspecified	Diagnosis			
155343005	Non-rheumatic mitral regurgitation	Diagnosis			
194726006	Mitral stenosis with insufficiency	Diagnosis			
156928009	Patent ductus arteriosus	Diagnosis			
204355003	Tricuspid atresia	Diagnosis			
204292000	(Persistent truncus arteriosus) or (Truncus	Diagnosis			
86299006	Tetralogy of Fallot	Diagnosis			

Table 6-2. SNOMED diagnoses for heart (left half) and lung sounds (right half). Terms from UMLS (2010).

Note that the ability to record and upload *compressed* audio from the Android phone (i.e. a clinician's voice notes) to OpenMRS was also added to the Sana infrastructure. The audio was recorded in 3GP format on the Android phone and sent to the MDS, where the file was converted to MP3 format using FFMPEG.[49] Then the MP3 audio file was uploaded to OpenMRS to be played back within the Flash media viewer, which supports the MP3 codec as mentioned above. This feature will be prepared for the second official release of the Sana source code.

[49] FFMPEG is an open-source library to convert audio and video between various file formats. http://www.ffmpeg.org/

6.3.7 System Documentation

In an effort to make the Sana platform easy to use by others across the globe with a working knowledge of English, detailed system installation and setup instructions were documented in a training manual for new users. The manual, which was written in January 2010, enabled new users to setup the Sana development environment to begin customizing the platform for their local deployment scenario. The documentation was then transferred to Sana's wiki site[50] in February 2010, which allowed frequent real-time updates and contributions from other developers. Below are the original contents of the Sana training manual that were added to the wiki:

1. Introduction
    - Walk Through a Demo of Sana
    - Download Source Code
2. Running Sana and Installing Tools for Development
    - Setting up the Basic Tools and Development Environment (Linux, Eclipse, Java, Android)
    - Detailed Instructions on How to Download the Source Code for Development (Android Phone Application, Sana Dispatch Server, OpenMRS modules)
    - Debugging with the Android, MDS, and OpenMRS Error Logs
3. Setting up Entire End-to-End Solution
    - Installing OpenMRS Server and Sana Dispatch Server
    - Configuring OpenMRS for Sana (Required Steps)
    - Learn about OpenMRS

[50] Sana wiki page (under its former name Moca): http://www.mocamobile.org/development/index.php?title=Main_Page

- Learn about Sana Dispatch Server

6.3.8 Security Considerations of Sana Platform

The HIPAA Security law, which was approved by the U.S. Congress in 2003, mandates that all U.S. healthcare institutions (such as insurance companies, health plans, or providers) follow certain rules and regulations to ensure the confidentiality and integrity of personal health information for all patients (U.S. Department of Health and Human Services 2003). In order for Sana to be deployed in the field in a manner compliant with the HIPAA Security Rule, policies must be implemented at the pilot site to fulfill 18 requirements that fall within three categories: administrative, physical, and technical safeguards (Weil 2004).

After an end-to-end system assessment of the Sana platform from a security standpoint, several areas for improvement were identified and preliminary technical solutions were made. This work was completed as part of a project on image encryption in the context of telemedicine for the MIT 6.857 Network and Computer Security course, which was performed in Spring 2010 (MIT OpenCourseWare 2003a) with classmate Adrian Dalca and technical help from RJ Ryan.

As part of the audit controls required by HIPAA, an institution must be able to provide, on request, written reports detailing access to data by all users to ensure that there have been no security violations. As a result, a logging system was prototyped in Sana by extending the Usage Statistics Module in OpenMRS (2010b). A new database table was created in OpenMRS to track the level of interaction with Sana cases through a usage ID number, user, patient involved, type of case, date and time of access, and

type of action performed on the data (viewed, created, diagnosed). See Figure 6-20 for a screenshot of the modified *Usage Statistics Module*.

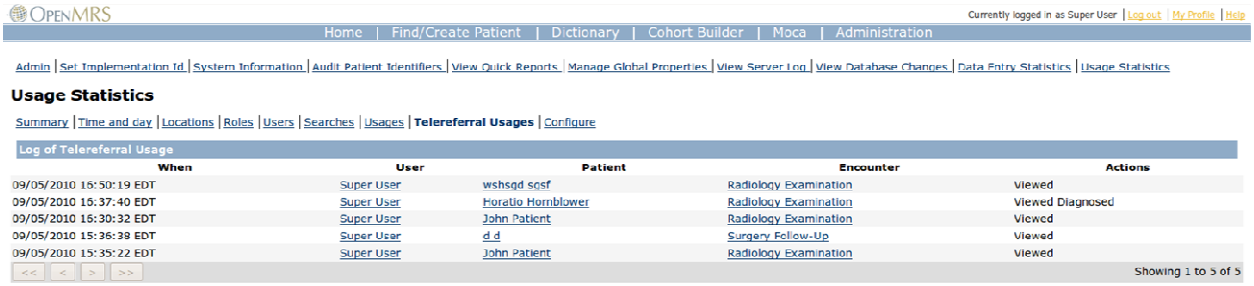


Figure 6-20. Usage Statistics module extended to track users activity on Sana patient cases.

Transmission security is one of the technical safeguards required by HIPAA, but communication with the MDS and OpenMRS servers occurs in an unencrypted manner. In order to enable SSL encryption[51] in Sana, the Apache server was configured to operate on port 443 for SSL and generate a self-signed certificate for testing purposes (a Versign[52] certificate would be obtained in the case of a real deployment), so that the MDS could be accessed securely. Then Apache was configured to proxy requests to Tomcat so that OpenMRS would also be accessed with a secure connection. In Figure 6-21, the use of an SSL connection to access OpenMRS are indicated by the yellow security padlock and URL containing 'https://'.

[51] Protocol to encrypt information in transactions made on the Internet by relying on the exchange of SSL certificates (which contain the public key of users) (James 2009).
[52] http://www.verisign.com/

Figure 6-21. Communication with OpenMRS server using SSL encryption.
See features highlighted in red.

The user is left with the decision to enable these security controls or not, and the instructions for configuring these settings are posted on a new page[53] in the Sana wiki. The page also includes instructions on how to password protect the MDS log and media files on the server to restrict access to network administrators only. Since some of the HIPAA mandates (such as facility security) are site-specific, the Sana wiki page also offers a discussion of issues regarding privacy and security for Sana implementers to consider as they design pilot studies. For example, methods of phone and computer security, system authentication, internal auditing, and training of personnel are discussed. (Note that the recommendations on the Sana wiki page are not comprehensive and do not necessarily make the system HIPAA compliant. End users of Sana should be aware that the software comes with no warranties or legal recourse).

[53] http://www.mocamobile.org/development/index.php?title=Security

6.3.9 Preparations for Sana Pilot in the Philippines

A six-week trip to the Philippines was made possible by the MIT Public Service Center to work with Sana's local community partner - the University of Philippines National Telehealth Center. During that time, preparations for a pilot of the Sana platform were made through meetings with key stakeholders from universities, telecommunication companies, hospitals, government agencies, and the PhilHealth[54] insurance corporation (see Figure 6-22).

Figure 6-22. Preparing for a deployment of Sana in the Philippines by speaking with all key stakeholders in the Philippines.

The focus of the trip was to evaluate the potential for a pilot in the Batanes Province, a rural location that is approximately a two-hour ($200) plane ride away from the capital city where specialist care is available. While Internet was unreliable in these rural areas, 3G connectivity was still available. Feasibility testing of Sana was therefore performed in Batanes (as a low-connectivity region), and cases with chest x-rays were successfully uploaded to the OpenMRS sever. Through the testing process, limitations and bugs in the system were discovered, which were prioritized to be fixed upon return to the U.S. in August 2009.

[54] http://www.philhealth.gov.ph/

Informational interviews and usability tests were also conducted to better understand the needs of the end-users and collect system requirements for the software. In the island province of Batanes, an X-ray machine was available, but there were no radiologists to interpret the films. Consequently, X-rays had to be sent to the Manila, the capital city of the Philippines, which meant a turn-around time of up to four weeks. There were also no doctors representing the other medical specialties either, as the province only had a limited number of general practitioners. In the case of serious health problems, the only option would be to fly to the Manila to seek care at the Philippine General Hospital. However, this was a decision that could deplete a family's savings and would be avoided at all costs. Based on anecdotal evidence from healthcare workers, midwives, doctors, and hospital staff in Batanes, it was clear that a telemedicine solution such as Sana could potentially be the vehicle on which to deliver specialized medical expertise to under-served rural areas similar to the Batanes.

The Sana platform could not be deployed sustainably though without financial support from local government units or PhilHealth national social insurance company. Considerable effort was made to encourage the adoption of eHealth technologies (including telemedicine) by PhilHealth. Meetings were held with their Chief Executive Officer, Chief Operating Officer, and members from the PhilHealth Informatics, Accreditation, Benefit Development, IT, Legal, and Fraud, Departments. In speaking with the employees of PhilHealth, a critical need was identified for the tracking of diseases and services used by individuals. These data could then be used by the corporation to achieve better disease management and develop more suitable benefit packages for underserved citizens. The paradigm (that investing in eHealth would

create better outcomes for all parties) is likely to take considerable time in order to be accepted and acted upon by those in charge of the corporation.

A complete report on the MIT Public Service Center fellowship work can be accessed from an entry[55] on the Sana blog. The report contains the following supplementary documents: Analysis of Philippine Healthcare System, Summary of eHealth Services of the National Telehealth Center, Matrix of eHealth and PhilHealth initiatives, Policy Proposal for PhilHealth, Sample Referral Forms for Workflow, Usability Testing of Sana, Usability Testing of OpenMRS Media Viewer, Technical Support Manual for Sana, and Summary of Workflow in the Batanes Province.

6.4 Discussion

In summary, the Sana platform has been extended to allow capture of heart and lung sounds from a mobile phone to a medical record system for review. CHWs, who may have not been previously trained on auscultation, can easily learn how to collect these sounds from patients by following step-by-step instructions on a mobile phone. Then instead of having to travel to a centralized location to manually transfer the data to a computer workstation, the audio recordings can be uploaded from any location in the field with cellular connectivity to a server. This reduces work, the number of lost and mislabeled records, and provides automated geo- and time-stamping. CHWs can also receive almost immediate responses[56] on the ausculatory recordings to diagnose their patients, receive feedback on recording technique, or receive requests to record sounds at more locations on the body. In this way, CHWs are educated to learn the technique of

[55] http://mocamobile.org/blog/?p=135
[56] Or at least before the patient or CHW have to part ways, thereby reducing the possibility that the user will not receive their diagnosis and follow the treatment recommendations.

auscultation by recognizing patterns in the sounds recorded from the patient with the corresponding diagnosis. Meanwhile, at the server side, a large database of annotated recordings can accumulate over time with the collective knowledge of many experts.

The functionality required at each step of the audio capture workflow has contributed in various improvements to the infrastructure of Sana. The Flash Media Viewer module served as foundational work on how complex patient data could be viewed in OpenMRS by Sana and other OpenMRS implementers. OpenMRS has demonstrated its continued support for this work by sponsoring another intern for Google Summer of Code 2010 to further develop the media viewer. Future versions of the media viewer will have the ability to display time series data (i.e. ECG or audio waveforms) and add text-based annotations on media files. Another goal is allow doctors to transcribe the audio and video file's contents within the media viewer for later reference without having to play the recording in its entirety again.

To complement the media viewer feature, the upload of numerically-coded vocabularies into OpenMRS provides a standardized approach to labeling the media files viewed in the medical record system. The assignment of diagnoses codes to patient cases allow for efficient data mining, which was demonstrated as a clear need in the Philippines by PhilHealth. For example, the database could potentially be searched for 'chronic obstructive pulmonary disorder' cases to identify all patients who have been diagnosed with the condition. Furthermore, their lung sounds, chest X-rays, or clinical symptoms could be extracted and analyzed for patterns to develop automated disease detection algorithms. Statistics can also be computed on the prevalence of different diseases in patient populations for the purpose of disease surveillance and public health

policy-making. In terms of improved patient care, the Sana encounter viewer can be extended to include an adjudication process whereby multiple experts can collaborate to decide on the final diagnosis for a patient (instead of relying on the first diagnosis from one specialist, which is likely to be inaccurate a significant proportion of the time (Neamatullah et al. 2008). It has been show that using three or more expert annotators to independently review medical data can lead to high diagnostic accuracy (at the 90-99% levels) (Neamatullah et al. 2008, Artstein and Poesio 2008, Bermingham and Smeaton 2009, Higgins 2007, Brants 2000, Gut and Bayerl 2004).

From a more technical standpoint, the system documentation on the Sana wiki site (described in Section 6.3.7) has helped to broaden the project's developer community with users across the globe in India and the Philippines now able to setup and begin customizing the platform. Other developers have also contributed to the documentation, while John Blakeney has created an automated installer for Sana based on the instructions. The security evaluation of the system and security solutions (from Section 6.3.8) will become necessary in Sana pilots that involve real patients in countries with governmental regulations on patient data confidentiality. For example, having a site-specific security policy should be an important part of training CHWs in any Sana pilot study.

Finally, the work in the Philippines served as preliminary groundwork in terms of building relationships with stakeholders, and gaining insight on how a mobile telemedicine platform could be used to improve clinical outcomes in a local context. Many of the findings emphasized that mobile health could enable providers to deliver previously unavailable specialized care to patients. For example, the ubiquity of mobile

phones even in the most remote communities of the Philippines was observed (i.e. in the Batanes, Pangasinan, and Batanges provinces), while supply chains for sophisticated medical devices were lacking (because of the extensive number of islands in the country). The development of medical diagnostic devices from locally available household materials would be one method to overcome poor supply chains in island provinces. The integration of diagnostic devices with mobile phones would allow transmission of data to specialists, which were desperately needed in areas where only general practitioners and volunteer healthcare workers resided such as Batanes. Much of the feedback gained from interacting with local healthcare workers enabled the Sana development team, over the following year, to better tailor the technical system to the needs of medical end-users.

Chapter 7
Conclusion

7.1 Thesis Contributions

This thesis describes a basic framework for leveraging mobile phones and the associated telecommunications infrastructure for the capture and remote annotation of heart and lung sounds, such that the signals can be processed intelligently to extract meaningful clinical information. First, gold standard databases of these sounds were used in the development and evaluation of algorithms for automated respiration and heart rate estimation. Pan and Tompkins' (1985) QRS peak detector for ECG signals was modified to successfully detect heart beats in heart sound signals (from an electronic stethoscope) with a sensitivity of 92.1% and a positive predictivity of 88.4%. Subsequently, on a database of lung sound signals, six respiration rate estimators were tested, and the TBC method had the lowest RMSPE with 4.34% and RMSE of 6.70 cpm.

Next, a physical system was developed to evaluate the feasibility of capturing heart and lung sounds with mobile phones. Based on thorough analysis of sound quality in a number of phones and hands-free kits, the Motorola Droid and iPhone hands-free kit were selected for use. (Other phones exhibited poor frequency responses in the frequency domain of clinical relevance.) Audio signals recorded on the surface of the body with the hands-free kit were deemed too weak and inaudible by expert listeners.

To remedy this issue, a low-cost homemade stethoscope was manufactured for acoustic amplification of the body's heart and lung sounds. The stethoscope device was created solely from household materials in the hopes that the device could be constructed and maintained both easily and cheaply in resource-poor communities.

The stethoscope attachment, hands-free kit, and Droid phone were then used in combination to capture heart and lung sounds from 50 test subjects. The result was the completion of a new database with 400 mobile phone recordings (100 of which were heart sounds recordings, while 300 were of lung sounds). The same algorithms, which had been tested on the standard databases, were applied to the field data. The heart rate estimation algorithm initially did not detect heart beats accurately on mobile phone recordings. However, the heart sound field data were split into test and training sets, while the training set was used to introduce several modifications to the heart rate estimation algorithm. The new heart rate estimator was ultimately able to yield an RMSPE of 1.55% and RMSE of 1.04 bpm on the test data consisting of 50 heart sound recordings. For respiratory rate estimation from lung sound field data, the TBC method also did not perform as well as it did on the standard lung sound database signals. Instead, a MEDIAN method (based on the median estimate from the TBC, MPT, and ARM1 methods) proved to have the lowest estimation error rates when compared to the other methods being tested, with an RMSPE of 8.38% and RMSE of 11.39 cpm. This result led to the conclusion that methods that worked best on standard databases may not perform similarly on data from the field.

Finally, the Sana telemedicine platform was also extended to allow the capture, upload, and labeling of heart and lung sounds as a way of providing remote specialists

with auscultation information on patients to make a timely and accurate diagnosis. Several key features that were contributed to the Sana platform included the Flash media viewer and integration of medical lexicons into OpenMRS, which open many opportunities for more sophisticated interaction with patient data in OpenMRS in the future.

7.2 Future Work

There are many future possibilities to explore, which can use the work in this thesis as a foundation for a system that can be adopted in resource poor locations, such as in rural communities of developing nations. Mechanical improvements can be made to the stethoscope attachment to optimize sound quality and amplification. The device should be verified for reproducibility in low-income environments and alternative materials that are acceptable substitutes should be identified. There should also be formal protocols on how to determine whether the stethoscope attachment has been manufactured accurately or not for use in the field. The chest piece on the device currently has a three-inch diameter, which may be a design concern when capturing sounds from an infant or neonate (whose bodies are too small compared to the device).

Improvements can also be made in the Sana phone application to make the workflow more convenient for the user (initially CHWs) when performing the auscultation procedures. One example would be to provide onscreen real-time feedback about the recording to help a CHW locate the optimal location on the body for recording. This feedback can occur as a real-time signal waveform displayed on the phone or a quantitative signal quality indicator. As evidenced in the low signal purity values of the

lung sounds in the database and field recordings, a superior method for computing signal quality of mobile phone heart and lung sounds is likely needed. The development of a new quality metric will be aided as more field data are collected and annotated. Automated diagnostic support can also be introduced to the platform by providing intelligent algorithms on the phone or OpenMRS interfaces, so that a diagnosis based on the input data could be deterministically suggested to users. Furthermore, in OpenMRS, the labeling of data using coded terms allows for future development of a wide range of reporting tools and disease management capabilities.

Ultimately, in order to prove that the capture of heart and lung sounds for remote diagnosis leads to improvements in patient clinical outcomes, a randomized clinical trial with a broad number of subjects is necessary. The test subjects from Chapter 5 were all adults over 18 years old, so a comprehensive study extending to young children and the elderly population would provide significant information on the algorithms' performance for all age groups. The trial would investigate whether heart and lung sounds can be used to diagnose specific conditions using the mobile phone-base system can provide comparable accuracy to that provided by a traditional stethoscope and a trained specialist (in-person). Validation of the framework described in this thesis would be a significant gain in being able to provide improved screening of respiratory and cardiac diseases among rural populations of developing nations.

Appendices

Appendix A: Original Heart Rate Estimation Method Results from Field Data

SUBJ ID	LOCATION ID	ACTUAL BPM	ESTIMATED BPM	PERCENTAGE ERROR
1	1	71.00	73.17	0.03
1	2	74.00	179.64	1.43
2	1	89.00	85.59	-0.04
2	2	89.00	225.56	1.53
3	1	74.00	72.99	-0.01
3	2	74.00	71.77	-0.03
4	1	77.00	75.57	-0.02
4	2	78.00	188.62	1.42
5	1	57.00	57.25	0.00
5	2	59.00	57.75	-0.02
6	1	75.00	75.38	0.01
6	2	79.00	200.00	1.53
7	1	82.00	82.42	0.01
7	2	84.00	82.42	-0.02
8	1	72.00	82.42	0.14
8	2	71.00	82.42	0.16
9	1	86.00	144.93	0.69
9	2	83.00	44.44	-0.46
10	1	79.00	133.93	0.70
10	2	88.00	79.37	-0.10
11	1	74.00	105.26	0.42
11	2	71.00	109.49	0.54
12	1	78.00	74.63	-0.04
12	2	75.00	74.63	NaN
13	1	49.00	87.72	0.79
13	2	51.00	45.87	-0.10
14	1	84.00	83.57	-0.01
14	1	88.00	88.50	0.01
15	1	63.00	168.88	1.68
15	1	89.00	189.87	1.13
16	1	73.00	128.21	0.76
16	2	82.00	126.58	0.54
17	1	91.00	89.02	-0.02
17	2	91.00	89.02	NaN
18	1	93.00	92.02	-0.01
18	2	96.00	228.88	1.38
19	1	90.00	240.00	1.67
19	2	99.00	177.51	0.79
20	1	89.00	85.71	-0.04
20	2	94.00	86.96	-0.07
21	1	69.00	65.93	-0.04
21	2	67.00	67.72	0.01
22	1	80.00	123.46	0.54
22	2	74.00	60.00	-0.19
23	1	75.00	174.42	1.33
23	2	70.00	Inf	Inf

24	**1**	89.00	Inf	NaN
24	**2**	91.00	91.19	0.00
25	**1**	49.00	71.60	0.46
25	**2**	59.00	63.29	0.07
26	**1**	82.00	63.29	NaN
26	**2**	85.00	229.02	1.69
27	**1**	82.00	84.03	0.02
27	**2**	82.00	78.95	-0.04
28	**1**	66.00	65.79	0.00
28	**2**	80.00	78.33	-0.02
29	**1**	69.00	72.12	0.05
29	**2**	62.00	72.12	NaN
30	**1**	53.00	56.71	0.07
30	**2**	59.00	56.71	NaN
31	**1**	93.00	139.53	0.50
31	**2**	84.00	124.24	0.48
32	**1**	83.00	81.30	-0.02
32	**2**	94.00	81.30	NaN
33	**1**	55.00	73.35	0.33
33	**1**	61.00	72.29	0.19
34	**1**	61.00	61.60	0.01
34	**2**	68.00	68.65	0.01
35	**1**	49.00	49.18	0.00
35	**2**	58.00	52.91	-0.09
36	**1**	83.00	65.79	-0.21
36	**2**	72.00	64.94	-0.10
37	**1**	88.00	47.39	-0.46
37	**2**	68.00	68.34	0.00
38	**1**	72.00	72.29	0.00
38	**2**	74.00	77.92	0.05
39	**1**	69.00	63.16	-0.08
39	**2**	60.00	62.63	0.04
40	**1**	79.00	74.08	-0.06
40	**2**	71.00	176.48	1.49
41	**1**	78.00	122.95	0.58
41	**2**	85.00	122.95	0.45
42	**1**	81.00	122.95	NaN
42	**2**	88.00	91.19	0.04
43	**1**	57.00	91.19	0.60
43	**2**	56.00	55.97	0.00
44	**1**	71.00	64.10	-0.10
44	**2**	63.00	61.10	-0.03
45	**1**	77.00	61.10	-0.21
45	**2**	73.00	61.10	NaN
46	**1**	61.00	61.10	NaN
46	**2**	58.00	56.07	-0.03
47	**1**	93.00	154.64	0.66
47	**2**	99.00	154.64	NaN
48	**1**	67.00	65.79	-0.02
48	**2**	68.00	65.79	-0.03
49	**1**	79.00	65.79	NaN
49	**2**	67.00	189.87	1.83
50	**1**	79.00	79.79	0.01
50	**2**	68.00	78.13	0.15

Appendix B: Pulse Oximeter Measurements from Field Data

SUBJ ID	AVERAGE HR AT LOCATION 1 (BPM)	AVERAGE HR AT LOCATION 2 (BPM)
1	71	74
2	89	89
3	74	74
4	77	78
5	57	59
6	75	79
7	82	84
8	72	71
9	86	83
10	79	88
11	74	71
12	78	75
13	49	51
14	88	84
15	63	89
16	73	82
17	91	91
18	93	96
19	90	99
20	89	94
21	69	67
22	80	74
23	75	70
24	89	91
25	49	59
26	82	85
27	82	82
28	66	80
29	69	62
30	53	59
31	93	84
32	83	94
33	61	55
34	61	68
35	49	58
36	83	72
37	88	68

38	72	74
39	69	60
40	79	71
41	78	85
42	81	88
43	57	56
44	71	63
45	77	73
46	61	58
47	93	99
48	67	68
49	79	67
50	79	74

Appendix C: Respiration Rate Estimation Results from Field Data

SUBJ ID	LOCATION ID	ACTUAL CPM	TBC	MPT	ARM1	ARM2	ARM3	MEDIAN	PURITY
1	1	14.00	0.71	0.70	-0.86	-0.86	-0.86	0.70	0.05
1	2	16.00	0.00	1.19	-0.88	-0.88	-0.88	0.00	0.03
1	3	13.00	0.54	1.69	-0.85	-0.85	-0.85	0.54	0.08
1	4	13.00	1.23	1.69	-0.85	-0.85	-0.85	1.23	0.04
1	5	14.00	0.86	1.45	-0.86	-0.86	-0.86	0.86	0.09
1	6	16.00	0.81	1.19	-0.88	-0.88	-0.88	0.81	0.05
2	1	14.00	1.21	1.50	1.12	-0.86	1.12	1.21	0.02
2	2	14.00	1.50	1.19	-0.86	-0.86	-0.86	1.19	0.01
2	3	15.00	0.13	1.33	-0.87	-0.87	-0.87	0.13	0.07
2	4	16.00	-0.19	1.19	0.92	-0.88	0.92	0.92	0.12
2	5	15.00	0.07	1.33	-0.87	-0.87	-0.87	0.07	0.11
2	6	14.00	0.71	0.41	-0.86	-0.86	-0.86	0.41	0.04
3	1	13.00	1.15	1.69	-0.85	-0.85	-0.85	1.15	0.04
3	2	18.00	-0.06	-0.22	-0.89	-0.89	-0.89	-0.22	0.10
3	3	11.00	0.36	2.18	-0.82	-0.82	-0.82	0.36	0.10
3	4	10.00	0.00	2.50	0.95	-0.80	0.95	0.95	0.18
3	5	14.00	1.50	1.50	-0.86	-0.86	-0.86	1.50	0.08
3	6	12.00	1.92	1.92	-0.83	-0.83	-0.83	1.92	0.06
4	1	24.00	0.46	0.46	-0.92	-0.92	-0.92	0.46	0.03
4	2	27.00	0.30	0.30	-0.93	-0.93	-0.93	0.30	0.02
4	3	24.00	0.46	0.46	-0.92	-0.92	-0.92	0.46	0.11
4	4	22.00	0.59	0.59	-0.91	-0.91	-0.91	0.59	0.05
4	5	26.00	0.31	0.35	-0.92	-0.92	-0.92	0.31	0.09
4	6	28.00	0.04	0.25	-0.93	-0.93	-0.93	0.04	0.05
5	1	14.00	0.93	0.20	-0.86	-0.86	-0.86	0.20	0.03
5	2	14.00	1.50	-0.50	-0.86	-0.86	-0.86	-0.50	0.01
5	3	13.00	0.31	1.09	-0.85	-0.85	-0.85	0.31	0.07
5	4	13.00	1.69	-0.69	-0.85	-0.85	-0.85	-0.69	0.02
5	5	12.00	1.92	1.02	-0.83	-0.83	-0.83	1.02	0.02
5	6	13.00	1.46	1.21	-0.85	-0.85	-0.85	1.21	0.05
6	1	15.00	1.33	0.34	-0.87	-0.87	-0.87	0.34	0.03
6	2	16.00	1.19	1.19	-0.88	-0.88	-0.88	1.19	0.02
6	3	15.00	1.33	-0.05	-0.87	-0.87	-0.87	-0.05	0.05
6	4	15.00	1.33	1.28	-0.87	-0.87	-0.87	1.28	0.02
6	5	14.00	1.21	0.63	-0.86	-0.86	-0.86	0.63	0.05
6	6	20.00	0.75	0.75	-0.90	-0.90	-0.90	0.75	0.01
7	1	18.00	0.44	0.94	-0.89	-0.89	-0.89	0.44	0.03
7	2	19.00	-0.21	-0.26	-0.89	-0.89	-0.89	-0.26	0.09
7	3	19.00	0.84	0.84	-0.89	-0.89	-0.89	0.84	0.01
7	4	20.00	0.75	0.63	-0.90	-0.90	-0.90	0.63	0.01
7	5	16.00	-0.13	0.39	0.79	-0.88	0.79	0.39	0.05
7	6	18.00	0.50	0.16	-0.89	-0.89	-0.89	0.16	0.03
8	1	12.00	1.92	1.92	-0.83	-0.83	-0.83	1.92	0.01
8	2	12.00	1.92	1.72	-0.83	-0.83	-0.83	1.72	0.01
8	3	13.00	0.00	1.45	1.22	-0.85	1.22	1.22	0.10
8	4	12.00	1.25	0.74	-0.83	-0.83	-0.83	0.74	0.03
8	5	13.00	0.15	0.08	-0.85	-0.85	-0.85	0.08	0.09
8	6	14.00	0.00	0.59	1.05	-0.86	1.05	0.59	0.05
9	1	16.00	1.19	0.21	-0.88	-0.88	-0.88	0.21	0.01
9	2	13.00	1.69	0.91	-0.85	-0.85	-0.85	0.91	0.01
9	3	16.00	0.38	-0.37	-0.88	-0.88	-0.88	-0.37	0.06
9	4	16.00	0.19	1.19	0.62	-0.88	0.62	0.62	0.04
9	5	14.00	0.64	-0.55	-0.86	-0.86	-0.86	-0.55	0.11

9	**6**	15.00	0.20	1.33	-0.87	-0.87	-0.87	0.20	0.09
10	**1**	14.00	0.43	1.50	1.07	-0.86	1.07	1.07	0.08
10	**2**	19.00	0.26	0.84	-0.89	-0.89	-0.89	0.26	0.03
10	**3**	16.00	0.19	1.19	-0.88	-0.88	-0.88	0.19	0.07
10	**4**	16.00	0.25	1.19	-0.88	-0.88	-0.88	0.25	0.05
10	**5**	15.00	0.87	1.33	-0.87	-0.87	-0.87	0.87	0.02
10	**6**	17.00	0.29	1.06	-0.88	-0.88	-0.88	0.29	0.04
11	**1**	18.00	0.94	0.94	-0.89	-0.89	-0.89	0.94	0.02
11	**2**	17.00	1.06	-0.06	-0.88	-0.88	-0.88	-0.06	0.04
11	**3**	18.00	0.94	0.30	-0.89	-0.89	-0.89	0.30	0.03
11	**4**	18.00	0.94	0.94	-0.89	-0.89	-0.89	0.94	0.01
11	**5**	17.00	0.53	0.66	-0.88	-0.88	-0.88	0.53	0.02
11	**6**	20.00	-0.40	0.75	0.50	-0.90	0.50	0.50	0.12
12	**1**	14.00	1.50	-0.41	-0.86	-0.86	-0.86	-0.41	0.03
12	**2**	14.00	0.21	0.11	-0.86	-0.86	-0.86	0.11	0.09
12	**3**	15.00	0.20	1.18	-0.87	-0.87	-0.87	0.20	0.03
12	**4**	14.00	1.50	1.50	-0.86	-0.86	-0.86	1.50	0.04
12	**5**	14.00	0.71	0.04	-0.86	-0.86	-0.86	0.04	0.07
12	**6**	14.00	0.36	-0.28	-0.86	-0.86	-0.86	-0.28	0.10
13	**1**	14.00	0.00	0.58	0.20	-0.86	0.20	0.20	0.06
13	**2**	14.00	1.50	0.77	-0.86	-0.86	-0.86	0.77	0.02
13	**3**	13.00	0.69	0.04	-0.85	-0.85	-0.85	0.04	0.08
13	**4**	15.00	1.27	1.33	-0.87	-0.87	-0.87	1.27	0.06
13	**5**	14.00	0.36	1.33	-0.86	-0.86	-0.86	-0.25	0.08
13	**6**	13.00	0.77	-0.57	-0.85	-0.85	-0.85	-0.57	0.06
14	**1**	15.00	0.60	-0.47	-0.87	-0.87	-0.87	-0.47	0.02
14	**2**	21.00	0.48	0.67	-0.90	-0.90	-0.90	0.48	0.03
14	**3**	14.00	-0.29	0.45	1.25	-0.86	1.25	0.45	0.07
14	**4**	16.00	1.00	-0.73	-0.88	-0.88	-0.88	-0.73	0.02
14	**5**	15.00	0.67	1.33	-0.87	-0.87	-0.87	0.67	0.03
14	**6**	15.00	-0.07	0.51	0.99	-0.87	0.99	0.51	0.17
15	**1**	11.00	0.09	2.18	1.60	-0.82	1.60	1.60	0.06
15	**2**	6.00	4.83	3.01	-0.67	-0.67	-0.67	3.01	0.02
15	**3**	14.00	0.50	0.68	-0.86	-0.86	-0.86	0.50	0.06
15	**4**	13.00	0.85	1.69	-0.85	-0.85	-0.85	0.85	0.05
15	**5**	15.00	0.07	1.33	1.18	-0.87	1.18	1.18	0.13
15	**6**	14.00	-0.14	1.50	0.98	-0.86	0.98	0.98	0.11
16	**1**	17.00	-0.88	0.02	-0.88	-0.88	-0.88	-0.88	0.11
16	**2**	17.00	1.06	0.14	-0.88	-0.88	-0.88	0.14	0.02
16	**3**	18.00	0.06	-0.07	-0.89	-0.89	-0.89	-0.07	0.04
16	**4**	20.00	0.45	-0.29	-0.90	-0.90	-0.90	-0.29	0.03
16	**5**	18.00	0.00	-0.41	-0.89	-0.89	-0.89	-0.41	0.08
16	**6**	20.00	0.20	-0.58	-0.90	-0.90	-0.90	-0.58	0.04
17	**1**	13.00	-0.08	0.98	0.64	-0.85	0.64	0.64	0.04
17	**2**	15.00	1.33	-0.19	-0.87	-0.87	-0.87	-0.19	0.02
17	**3**	12.00	1.92	0.01	-0.83	-0.83	-0.83	0.01	0.02
17	**4**	13.00	0.31	1.69	-0.85	-0.85	-0.85	0.31	0.06
17	**5**	12.00	0.17	1.92	0.66	-0.83	0.66	0.66	0.11
17	**6**	11.00	0.09	1.34	0.94	-0.82	0.94	0.94	0.04
18	**1**	13.00	1.69	0.07	-0.85	-0.85	-0.85	0.07	0.01
18	**2**	9.00	2.89	-0.55	-0.78	-0.78	-0.78	-0.55	0.01
18	**3**	12.00	0.83	0.51	1.27	-0.83	1.27	0.83	0.03
18	**4**	13.00	1.69	0.69	-0.85	-0.85	-0.85	0.69	0.01
18	**5**	13.00	1.23	0.22	-0.85	-0.85	-0.85	0.22	0.02
18	**6**	12.00	0.42	-0.12	-0.83	-0.83	-0.83	-0.12	0.04
19	**1**	24.00	0.42	-0.66	-0.92	-0.92	-0.92	-0.66	0.03
19	**2**	24.00	0.33	0.46	-0.92	-0.92	-0.92	0.33	0.02
19	**3**	24.00	-0.21	-0.16	-0.92	-0.92	-0.92	-0.21	0.08
19	**4**	22.00	0.14	-0.83	-0.91	-0.91	-0.91	-0.83	0.09
19	**5**	22.00	0.55	-0.91	-0.91	-0.91	-0.91	-0.91	0.05
19	**6**	23.00	0.17	-0.38	-0.91	-0.91	-0.91	-0.38	0.06
20	**1**	20.00	0.70	0.75	-0.90	-0.90	-0.90	0.70	0.04

20	**2**	20.00	-0.10	-0.44	-0.90	-0.90	-0.90	-0.44	0.05
20	**3**	19.00	-0.16	-0.76	-0.89	-0.89	-0.89	-0.76	0.04
20	**4**	18.00	0.94	-0.32	-0.89	-0.89	-0.89	-0.32	0.03
20	**5**	20.00	0.10	0.75	-0.90	-0.90	-0.90	0.10	0.12
20	**6**	18.00	0.94	0.94	-0.89	-0.89	-0.89	0.94	0.05
21	**1**	16.00	0.50	1.19	-0.88	-0.88	-0.88	0.50	0.06
21	**2**	17.00	1.06	0.32	-0.88	-0.88	-0.88	0.32	0.02
21	**3**	16.00	0.25	0.36	1.00	-0.88	1.00	0.36	0.09
21	**4**	18.00	0.11	0.94	-0.89	-0.89	-0.89	0.11	0.05
21	**5**	16.00	0.50	-0.65	-0.88	-0.88	-0.88	-0.65	0.04
21	**6**	14.00	1.50	0.80	-0.86	-0.86	-0.86	0.80	0.01
22	**1**	12.00	0.75	1.92	1.54	-0.83	1.54	1.54	0.05
22	**2**	12.00	1.42	0.59	-0.83	-0.83	-0.83	0.59	0.02
22	**3**	14.00	0.93	-0.01	-0.86	-0.86	-0.86	-0.01	0.02
22	**4**	14.00	-0.21	0.64	0.89	-0.86	0.89	0.64	0.05
22	**5**	15.00	1.33	0.92	-0.87	-0.87	-0.87	0.92	0.03
22	**6**	15.00	1.33	0.24	-0.87	-0.87	-0.87	0.24	0.02
23	**1**	11.00	0.64	0.53	-0.82	-0.82	-0.82	0.53	0.03
23	**2**	10.00	1.60	0.76	-0.80	-0.80	-0.80	0.76	0.02
23	**3**	11.00	-0.09	0.28	1.64	-0.82	1.64	0.28	0.07
23	**4**	12.00	1.92	1.92	-0.83	-0.83	-0.83	1.92	0.02
23	**5**	12.00	1.92	1.92	-0.83	-0.83	-0.83	1.92	0.02
23	**6**	10.00	2.50	2.50	-0.80	-0.80	-0.80	2.50	0.01
24	**1**	18.00	-0.33	-0.47	0.61	-0.89	0.61	-0.33	0.06
24	**2**	20.00	0.20	-0.45	-0.90	-0.90	-0.90	-0.45	0.04
24	**3**	19.00	0.00	0.24	-0.89	-0.89	-0.89	0.00	0.07
24	**4**	19.00	0.26	0.84	-0.89	-0.89	-0.89	0.26	0.07
24	**5**	20.00	0.60	-0.41	-0.90	-0.90	-0.90	-0.41	0.04
24	**6**	22.00	0.59	0.59	-0.91	-0.91	-0.91	0.59	0.03
25	**1**	16.00	0.19	1.19	0.76	-0.88	0.76	0.76	0.05
25	**2**	15.00	1.20	0.27	-0.87	-0.87	-0.87	0.27	0.02
25	**3**	15.00	-0.47	1.10	0.54	-0.87	0.54	0.54	0.10
25	**4**	14.00	-0.36	1.50	0.67	-0.86	0.67	0.67	0.08
25	**5**	14.00	0.71	-0.35	0.80	-0.86	0.80	0.71	0.03
25	**6**	14.00	0.14	1.50	0.91	-0.86	0.91	0.91	0.06
26	**1**	17.00	-0.41	1.06	0.62	-0.88	0.62	0.62	0.10
26	**2**	12.00	1.92	1.92	-0.83	-0.83	-0.83	1.92	0.01
26	**3**	16.00	0.00	0.76	-0.88	-0.88	-0.88	0.00	0.07
26	**4**	13.00	0.15	-0.11	-0.85	-0.85	-0.85	-0.11	0.03
26	**5**	17.00	-0.41	-0.32	0.46	-0.88	0.46	-0.32	0.06
26	**6**	13.00	-0.15	1.69	0.89	-0.85	0.89	0.89	0.05
27	**1**	21.00	0.67	-0.34	-0.90	-0.90	-0.90	-0.34	0.02
27	**2**	22.00	0.00	-0.10	-0.91	-0.91	-0.91	-0.10	0.12
27	**3**	24.00	0.38	0.46	-0.92	-0.92	-0.92	0.38	0.04
27	**4**	23.00	0.52	-0.33	-0.91	-0.91	-0.91	-0.33	0.03
27	**5**	26.00	0.12	-0.09	-0.92	-0.92	-0.92	-0.09	0.05
27	**6**	26.00	-0.73	0.35	-0.45	-0.92	-0.45	-0.45	0.12
28	**1**	14.00	0.43	1.50	0.90	-0.86	0.90	0.90	0.08
28	**2**	15.00	1.00	-0.41	-0.87	-0.87	-0.87	-0.41	0.06
28	**3**	14.00	0.79	1.50	-0.86	-0.86	-0.86	0.79	0.10
28	**4**	14.00	0.29	1.50	-0.86	-0.86	-0.86	0.29	0.08
28	**5**	14.00	0.93	1.50	-0.86	-0.86	-0.86	0.93	0.11
28	**6**	15.00	0.13	-0.58	1.04	-0.87	1.04	0.13	0.12
29	**1**	12.00	1.92	0.87	-0.83	-0.83	-0.83	0.87	0.01
29	**2**	11.00	2.18	0.83	-0.82	-0.82	-0.82	0.83	0.01
29	**3**	11.00	0.09	0.32	1.25	-0.82	1.25	0.32	0.13
29	**4**	14.00	0.36	1.50	-0.86	-0.86	-0.86	0.36	0.04
29	**5**	12.00	-0.17	1.92	0.48	-0.83	0.48	0.48	0.14
29	**6**	12.00	0.33	-0.35	1.16	-0.83	1.16	0.33	0.07
30	**1**	17.00	0.12	-0.34	-0.88	-0.88	-0.88	-0.34	0.03
30	**2**	18.00	0.50	0.94	-0.89	-0.89	-0.89	0.50	0.03
30	**3**	19.00	0.84	0.84	-0.89	-0.89	-0.89	0.84	0.05

30	**4**	17.00	1.06	1.06	-0.88	-0.88	-0.88	1.06	0.03
30	**5**	20.00	-0.25	0.47	0.56	-0.90	0.56	0.47	0.09
30	**6**	17.00	1.06	0.37	-0.88	-0.88	-0.88	0.37	0.03
31	**1**	15.00	0.73	-0.41	-0.87	-0.87	-0.87	-0.41	0.01
31	**2**	16.00	0.19	-0.79	-0.88	-0.88	-0.88	-0.79	0.03
31	**3**	14.00	-0.36	1.00	0.53	-0.86	0.53	0.53	0.21
31	**4**	15.00	0.67	0.59	-0.87	-0.87	-0.87	0.59	0.04
31	**5**	16.00	0.94	-0.59	-0.88	-0.88	-0.88	-0.59	0.04
31	**6**	14.00	1.07	0.47	-0.86	-0.86	-0.86	0.47	0.08
32	**1**	14.00	0.86	0.72	-0.86	-0.86	-0.86	0.72	0.07
32	**2**	14.00	1.50	1.50	-0.86	-0.86	-0.86	1.50	0.03
32	**3**	16.00	0.25	-0.13	-0.88	-0.88	-0.88	-0.13	0.04
32	**4**	15.00	0.33	0.59	-0.87	-0.87	-0.87	0.33	0.09
32	**5**	18.00	0.11	0.21	-0.89	-0.89	-0.89	0.11	0.03
32	**6**	15.00	1.33	-0.18	-0.87	-0.87	-0.87	-0.18	0.02
33	**1**	13.00	1.69	1.69	-0.85	-0.85	-0.85	1.69	0.01
33	**2**	11.00	0.27	1.15	-0.82	-0.82	-0.82	0.27	0.04
33	**3**	12.00	1.00	0.59	-0.83	-0.83	-0.83	0.59	0.02
33	**4**	13.00	1.69	0.04	-0.85	-0.85	-0.85	0.04	0.03
33	**5**	14.00	1.50	0.93	-0.86	-0.86	-0.86	0.93	0.01
33	**6**	12.00	1.00	0.59	-0.83	-0.83	-0.83	0.59	0.02
34	**1**	15.00	0.47	1.33	-0.87	-0.87	-0.87	0.47	0.07
34	**2**	16.00	0.25	1.19	1.09	-0.88	1.09	1.09	0.05
34	**3**	16.00	0.00	-0.35	-0.88	-0.88	-0.88	-0.35	0.11
34	**4**	15.00	0.87	-0.10	-0.87	-0.87	-0.87	-0.10	0.05
34	**5**	17.00	0.12	1.06	-0.88	-0.88	-0.88	0.12	0.12
34	**6**	16.00	1.19	0.44	-0.88	-0.88	-0.88	0.44	0.02
35	**1**	13.00	0.69	1.69	0.26	-0.85	0.26	0.69	0.03
35	**2**	15.00	1.33	1.33	-0.87	-0.87	-0.87	1.33	0.02
35	**3**	14.00	0.00	0.79	1.28	-0.86	1.28	0.79	0.06
35	**4**	14.00	0.86	1.21	-0.86	-0.86	-0.86	0.86	0.05
35	**5**	14.00	0.14	1.50	-0.86	-0.86	-0.86	0.14	0.13
35	**6**	14.00	0.64	1.50	-0.86	-0.86	-0.86	0.64	0.04
36	**1**	10.00	-0.50	2.50	0.25	-0.80	0.25	0.25	0.13
36	**2**	10.00	2.50	2.50	-0.80	-0.80	-0.80	2.50	0.02
36	**3**	8.00	3.38	0.65	-0.75	-0.75	-0.75	0.65	0.01
36	**4**	10.00	2.50	-0.15	-0.80	-0.80	-0.80	-0.15	0.02
36	**5**	11.00	2.18	1.39	-0.82	-0.82	-0.82	1.39	0.11
36	**6**	12.00	1.92	1.92	-0.83	-0.83	-0.83	1.92	0.02
37	**1**	12.00	0.42	0.69	-0.83	-0.83	-0.83	0.42	0.08
37	**2**	13.00	1.69	0.65	-0.85	-0.85	-0.85	0.65	0.01
37	**3**	13.00	1.08	1.69	-0.85	-0.85	-0.85	1.08	0.04
37	**4**	13.00	0.15	-0.05	-0.85	-0.85	-0.85	-0.05	0.02
37	**5**	14.00	-0.43	1.30	0.56	-0.86	0.56	0.56	0.11
37	**6**	13.00	0.54	0.51	-0.85	-0.85	-0.85	0.51	0.07
38	**1**	17.00	1.06	0.35	-0.88	-0.88	-0.88	0.35	0.02
38	**2**	18.00	0.94	0.39	-0.89	-0.89	-0.89	0.39	0.01
38	**3**	18.00	-0.50	-0.22	0.64	-0.89	0.64	-0.22	0.11
38	**4**	19.00	0.21	0.16	-0.89	-0.89	-0.89	0.16	0.06
38	**5**	20.00	0.05	-0.35	-0.90	-0.90	-0.90	-0.35	0.07
38	**6**	17.00	0.65	0.68	-0.88	-0.88	-0.88	0.65	0.12
39	**1**	9.00	2.33	-0.49	-0.78	-0.78	-0.78	-0.49	0.02
39	**2**	9.00	2.89	1.26	-0.78	-0.78	-0.78	1.26	0.02
39	**3**	10.00	2.50	1.67	-0.80	-0.80	-0.80	1.67	0.02
39	**4**	9.00	2.89	2.89	-0.78	-0.78	-0.78	2.89	0.02
39	**5**	8.00	1.50	3.38	-0.75	-0.75	-0.75	1.50	0.05
39	**6**	10.00	2.00	1.38	-0.80	-0.80	-0.80	1.38	0.03
40	**1**	13.00	0.69	1.69	-0.85	-0.85	-0.85	0.69	0.08
40	**2**	12.00	0.00	1.35	1.47	-0.83	1.47	1.35	0.08
40	**3**	12.00	0.75	1.71	1.67	-0.83	1.67	1.67	0.16
40	**4**	11.00	0.64	2.18	-0.82	-0.82	-0.82	0.64	0.18
40	**5**	12.00	0.58	1.15	-0.83	-0.83	-0.83	0.58	0.07

40	**6**	13.00	0.85	1.28	-0.85	-0.85	-0.85	0.85	0.05
41	**1**	11.00	1.64	2.18	-0.82	-0.82	-0.82	1.64	0.04
41	**2**	13.00	1.69	1.69	-0.85	-0.85	-0.85	1.69	0.02
41	**3**	11.00	2.18	0.27	-0.82	-0.82	-0.82	0.27	0.02
41	**4**	10.00	2.50	2.50	-0.80	-0.80	-0.80	2.50	0.01
41	**5**	11.00	1.55	2.18	-0.82	-0.82	-0.82	1.55	0.03
41	**6**	11.00	0.36	2.18	-0.82	-0.82	-0.82	0.36	0.03
42	**1**	15.00	1.33	1.33	-0.87	-0.87	-0.87	1.33	0.01
42	**2**	8.00	-0.13	1.05	1.23	-0.75	1.23	1.05	0.09
42	**3**	14.00	1.50	1.50	-0.86	-0.86	-0.86	1.50	0.03
42	**4**	11.00	2.18	1.59	-0.82	-0.82	-0.82	1.59	0.02
42	**5**	11.00	0.18	0.20	-0.82	-0.82	-0.82	0.18	0.08
42	**6**	12.00	0.42	1.92	1.50	-0.83	1.50	1.50	0.05
43	**1**	16.00	-0.38	1.19	0.72	-0.88	0.72	0.72	0.11
43	**2**	15.00	0.07	0.51	-0.87	-0.87	-0.87	0.07	0.04
43	**3**	16.00	0.88	0.28	-0.88	-0.88	-0.88	0.28	0.03
43	**4**	16.00	1.19	0.20	-0.88	-0.88	-0.88	0.20	0.02
43	**5**	14.00	0.57	1.50	-0.86	-0.86	-0.86	0.57	0.05
43	**6**	15.00	-0.07	1.33	1.15	-0.87	1.15	1.15	0.06
44	**1**	13.00	0.62	-0.30	-0.85	-0.85	-0.85	-0.30	0.03
44	**2**	14.00	1.50	1.50	-0.86	-0.86	-0.86	1.50	0.02
44	**3**	14.00	0.50	1.50	-0.86	-0.86	-0.86	0.50	0.03
44	**4**	14.00	0.29	1.50	-0.86	-0.86	-0.86	-0.29	0.08
44	**5**	16.00	1.19	0.52	-0.88	-0.88	-0.88	0.52	0.02
44	**6**	24.00	0.46	0.12	-0.92	-0.92	-0.92	0.12	0.05
45	**1**	22.00	0.45	0.16	-0.91	-0.91	-0.91	0.16	0.04
45	**2**	26.00	0.35	-0.45	-0.92	-0.92	-0.92	-0.45	0.03
45	**3**	27.00	-0.37	0.30	0.12	-0.93	0.12	0.12	0.07
45	**4**	22.00	0.14	0.20	-0.91	-0.91	-0.91	0.14	0.02
45	**5**	27.00	-0.37	-0.15	-0.93	-0.93	-0.93	-0.37	0.04
45	**6**	16.00	1.19	-0.05	-0.88	-0.88	-0.88	-0.05	0.02
46	**1**	15.00	1.33	1.29	-0.87	-0.87	-0.87	1.29	0.01
46	**2**	16.00	1.19	1.19	-0.88	-0.88	-0.88	1.19	0.01
46	**3**	15.00	1.13	-0.48	-0.87	-0.87	-0.87	-0.48	0.07
46	**4**	15.00	1.33	1.33	-0.87	-0.87	-0.87	1.33	0.01
46	**5**	15.00	0.80	1.12	-0.87	-0.87	-0.87	0.80	0.04
46	**6**	14.00	1.29	-0.13	-0.86	-0.86	-0.86	-0.13	0.04
47	**1**	19.00	-0.26	0.84	-0.89	-0.89	-0.89	-0.26	0.06
47	**2**	17.00	0.00	1.06	-0.88	-0.88	-0.88	0.00	0.08
47	**3**	19.00	-0.05	-0.04	-0.89	-0.89	-0.89	-0.05	0.08
47	**4**	16.00	0.44	1.19	-0.88	-0.88	-0.88	0.44	0.08
47	**5**	16.00	0.13	-0.42	-0.88	-0.88	-0.88	-0.42	0.10
47	**6**	18.00	-0.06	0.94	0.72	-0.89	0.72	0.72	0.12
48	**1**	40.00	-0.13	-0.78	-0.95	-0.95	-0.95	-0.78	0.03
48	**2**	40.00	-0.13	-0.68	-0.95	-0.95	-0.95	-0.68	0.04
48	**3**	40.00	-0.13	-0.13	-0.95	-0.95	-0.95	-0.13	0.04
48	**4**	40.00	-0.13	-0.13	-0.95	-0.95	-0.95	-0.13	0.07
48	**5**	30.00	0.17	-0.48	-0.93	-0.93	-0.93	-0.48	0.06
48	**6**	38.00	-0.08	-0.85	-0.95	-0.95	-0.95	-0.85	0.05
49	**1**	20.00	0.65	0.57	-0.90	-0.90	-0.90	0.57	0.02
49	**2**	20.00	0.50	-0.34	-0.90	-0.90	-0.90	-0.34	0.04
49	**3**	23.00	-0.39	0.42	-0.29	-0.91	-0.29	-0.29	0.09
49	**4**	22.00	0.59	-0.70	-0.91	-0.91	-0.91	-0.70	0.02
49	**5**	19.00	-0.42	0.84	-0.89	-0.89	-0.89	-0.42	0.06
49	**6**	22.00	-0.36	0.59	-0.91	-0.91	-0.91	-0.36	0.05
50	**1**	10.00	0.00	0.45	1.31	-0.80	1.31	0.45	0.11
50	**2**	10.00	2.50	-0.03	-0.80	-0.80	-0.80	-0.03	0.01
50	**3**	9.00	0.00	0.68	1.02	-0.78	1.02	0.68	0.13
50	**4**	10.00	1.60	1.21	-0.80	-0.80	-0.80	1.21	0.02
50	**5**	10.00	0.60	2.50	1.60	-0.80	1.60	1.60	0.07
50	**6**	9.00	1.78	1.65	-0.78	-0.78	-0.78	1.65	0.02

Bibliography

3M. *History of 3M Littmann Stethoscopes at a glance.* [Online] Littmann Stethoscopes Product Information. Available at: http://solutions.3m.com/wps/portal/3M/en_US/Littmann/stethoscope/products/history/ [Accessed 11 March 2010].

Acero, A., 1993. *Acoustical and environmental robustness in automatic speech recognition*. p. 5, Norwell, MA: Kluwer Academic Publishers.

Advanced Life Support Group, 2004. *Advanced Pediatric Life Support: The Practical Approach*. 4th ed., WileyBlackwell.

Ait-Khaled, N., Enarson, D., Bousquet, J., 2001. Chronic respiratory diseases in developing countries: the burden and strategies for prevention and management. *Bulletin of the World Health Organization*, 79:10, Geneva. (doi: 10.1590/S0042-96862001001000011)

American Heart Association, 2006. *Pediatric Advanced Life Support Provider Manual.* American Heart Association.

American Heart Association, 2010. *Cardiomyopathy*. [Online] Available at: http://www.americanheart.org/presenter.jhtml?identifier=4468/ [Accessed 2 July 2010].

American Thoracic Society, 1995. Standardization of spirometry. 1994 update. *Am J Respir Crit Care Med*, pp.1107–1136.

Anand, M., 2006. *Electronic Instruments and Instrumentation Technology.* 4th ed, Prentice Hall of India, New Delhi, p. 297.

Anders, H., 1907. *Physical Diagnosis with Case Examples of the Inductive Method.* D. Appelton and Company, New York, p. 132.

Anders, H. and Boston, L., 1911. *A text-book of medical diagnosis.* W.B. Saunders Company, Philadelphia, p. 61.

Artstein, R. and Poesio, M., 2008. Survey article: Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34:4 (5 July), pp. 555–596.

Ashby, H., 1905. *The Diseases of Children.* 5th ed., Longmans, Green, and Co., London, p. 345.

Ashley, E. and Niebauer, J., 2004, *Cardiology Explained*, Remedica Explained Series, Remedica, London.

Association for the Advancement of Medical Instrumentation 2003, *Testing and reporting performance results of cardiac rhythm and ST-segment measurement algorithms,* ANSI/AAMI EC57:1998/(R)2003, American National Standards Institute, Arlington.

Association for the Advancement of Medical Instrumentation 2007, *Medical electrical equipment – Part 2-47: Particular requirements for the safety, including essential performance, of ambulatory electrocardiographic systems,* ANSI/AAMI EC38:2007, American National Standards Institute, Arlington.

BabyBeat, 2009. *BabyBeat Fetal Dopplers* [Online] (Updated 2009) Available at: http://www.babybeat.com/ [Accessed 13 September 2009].

Bates, B., 1995. *A Guide to Physical Examination and History Making*. Lippincott, Philadelphia.

Bellos, A., Mulholland, K., O'Brien, K., Qazi, S., Gayer, M., Checchi, F., 2010. The burden of acute respiratory infections in crisis-affected populations: a systematic review. *Conflict and Health*, World Health Organization, 4:3.

Bentley, P., 2009. *iStethoscope Pro support page.* The World of Peter J. Bentley. [Online] (Updated 30 June 2009) Available at: http://apps.peterjbentley.com/Blog/?e=31790&d=06/30/2009&s=iStethoscope%20Pro%20support%20page [Accessed 15 September 2009].

Bermingham, A. and Smeaton, A., 2009. A Study of Inter-Annotator Agreement for Opinion Retrieval. *In Proceedings of SIGIR (2009),* July 19–2 , Boston, Massachusetts. (doi: 10.1145/1571941.1572127)

Bill and Melinda Gates Foundation, 2010. *Discovery: Strategy Overview.* Global Health Program. [Report] Available at: http://www.gatesfoundation.org/global-health/Documents/Discovery-strategy.pdf [Accessed 28 April 2010].

Blaya, J., 2008. *Developing, Implementing, and Evaluating Tuberculosis Laboratory Information Systems for Resource-Poor Settings*, PhD thesis, Massachusetts Institute of Technology, Cambridge. (uri: http://hdl.handle.net/1721.1/47852)

Blaya, J., Cohen, T., Rodriguez, P., Kim, J., and Fraser, H., 2009. Personal digital assistants to collect tuberculosis bacteriology data in Peru reduce delays, errors, and workload, and are acceptable to users: cluster randomized controlled trial. *Int J Infect Dis.*, 13:3 (May), pp. 410-418.

Blaya, J. and Fraser, H., 2006. *Development, implementation and preliminary study of a PDA-based bacteriology collection system,* American Medical Informatics Association Symposium Proceedings, 2006, pp. 41-45.

Blood, D. C., 2007. *Saunders comprehensive veterinary dictionary*. 3rd edn, Elsevier, Edinburgh, UK.

Bowden, V., Greenberg, C., 2009. *Children and Their Families: The Continuum of Care.* 2nd ed., Wolters Kluwer, Lippincott Williams & Wilkins, Philadelphia, p. 658.

Brants, T., 2000. Inter-annotator agreement for a German newspaper corpus. *In Proceedings of the Second International Conference on Language Resources and Evaluation* (LREC 2000), May 31- June 2, Athens, Greece.

Brown, E., Leung, T., Collis, W., 2008. *Heart sounds made easy.* 2nd ed., Elsevier Health Sciences, Philadelphia.

Burke, T., 2008. *Biomedical Engineering Assignment-ECG Data*. School of Electrical Engineering at Dublin Institute of Technology [Online] (Updated 2 April 2008) Available at: http://eleceng.dit.ie/tburke/biomed/ecg.png

Cabot, R., 1912. *Physical diagnosis.* 5th ed., William Wood and Company, New York.
Caldwell, J., 2001. Population health in transition. *Bulletin of the World Health Organization*, 79:2, Geneva. (doi: 10.1590/S0042-96862001000200011)

Carter, B. and Mancini, R., 2009. *Op Amps for Everyone.* 3rd ed, Elsevier, Oxford, pp. 213-214.

Centers for Disease Control and Prevention, 2010. *Healthy Weight: Assessing your weight: BMI: About Adult BMI.* [Online] (Updated 27 July 2009) Available at: http://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html [Accessed 20 April 2010].

Chandrasekhar, A., 1999. *Auscultation of Lungs: Back View*. Loyola University Medical Education Network [Online] (Updated 20 December 1999) Available at: http://www.meddean.luc.edu/lumen/MedEd/medicine/pulmonar/images/peimage/backview.jpg [Accessed 6 May 2010].

Chang, G. C., Cheng, Y. P., 2008. Investigation of Noise Effect on Lung Sound Recognition. In *IEEE 7th International Conference on Machine Learning and Cybernetics 2008*, 3 (12-15 July), Kunming, China, pp. 1298-1301. (doi: 10.1109/ICMLC.2008.4620605)

Chen T., Kuan, K., Celi, L. and Clifford G.D., 2010. Intelligent Heartsound Diagnostics on a Cellphone using a Hands-free Kit, AAAI Spring Symposium on Artificial Intelligence for Development (AI-D), Stanford University, March 22-24, 2010.

Cisco Systems, 2007. *Defining Analog Voice* [Report 8628] (Updated 17 February 2007) Gateway Protocols. Available at: http://www.cisco.com/application/pdf/paws/8628/define_analog_voice.pdf [Accessed 24 April 2010].

Clifford, G. D., 2002. *Signal Processing Methods For Heart Rate Variability,* Ph.D. Thesis, University of Oxford, December 2002.

Clifford, G.D., Blaya, J.A., Hall-Clifford, R., and Fraser, H.S.F., 2008. Medical Information Systems: A foundation for healthcare technologies in developing countries, BioMedical Engineering OnLine. 7:18 (11 June).

Clifford, G.D., Long, W.J., Moody, G.B., and Szolovits, P., 2009. Robust parameter extraction for decision support using multimodal intensive care data, Proc. Royal Soc. Phil. Trans. A, January 28, 2009, 367:1887, pp. 411-429. (doi: 10.1098/rsta.2008.0157)

Cohen, A. and Landsberg, D., 1990. Signal Processing Methods for Upper Airway and Pulmonary Dysfunction Diagnosis. *IEEE Engineering in Medicine and Biology Magazine*, 9(1), pp. 72-75. (doi: 10.1109/51.62912)

Contec Medical Systems, 2010. *Contec Medical Systems Co., Ltd.* [Online] Available at: http://www.contecmedical.net/ [Accessed 20 April 2010].

Davis, G., Seward, E., and Marcy, T., 1999. *Medical management of pulmonary diseases.* Marcel Dekker, New York, pp. 114-120.

Day, R., Paul, P., and Williams, B., 2009. *Brunner and Suddarth's Textbook of Canadian Medical-Surgical Nursing*, Unit 6: Cardiovascular, Circulatory, and Hematologic Function, Lippincott Williams & Wilkins, pp. 752-753.

Diederich, J., 2008. *Rule extraction from support vector machines.* Springer-Verlag, Heidelberg, p. 208.

Douglas, G., Nicol, E., and Robertson, C., 2009. *Macleod's clinical examination.* Churchill Livingstone Elsevier, U.S., p. 128.

Dreeben, O., 2008. *Physical Therapy Clinical Handbook for PTA's*. Jones and Bartlett Publishers , Sudbury, p. 41.

Earis, J., 1992. Lung Sounds. *Thorax International Journal of Respiratory Medicine*, 47, pp. 671-672. (doi: 10.1136/thx.47.9.671)

Eng, T., 2001. *The eHealth landscape: a terrain map of emerging information and communication technologies in health and healthcare.* Robert Wood Johnson Foundation, Princeton, NJ.

EpiHandy, 2008. EpiHandyMobile. (Updated 5 December 2008) Available at: http://www.epihandy.org/index.php/EpiHandyMobile [Accessed 15 July 2010].

Fenton, T., Pasterkamp, H., Tal, A., and Chernick, V., 1985. Automated Spectral Characterization of Wheezing in Asthmatic Children. *IEEE Transactions on Biomedical Engineering,* BME-32:1 (January), pp. 50-55. (doi: 10.1109/TBME.1985.325616)

Fleming, S. and Tarassenko, L., 2007. A Comparison of Signal Processing Techniques for the Extraction of Breathing Rate from the Photoplethysmogram. *Proceedings of World Academy of Science, Engineering and Technology*, 30, pp. 276-280.

Fleming, S., Thompson, M., Stevens, R., Heneghan, C., Pluddemann, A., Maconochie, I., Tarassenko, L., and Mant, D., 2010. Normal ranges of heart rate and respiratory rate in children from birth to 18 years: a systematic review of observational studies. *The Lancet*, (in press).

Fortner, P., 2009. *Mobile Phones Drive Health IT Innovation in Developing Countries.* iHealthBeat. [Online] (Updated 10 August 2009) Available at: http://www.ihealthbeat.org/Features/2009/Mobile-Phones-Driving-Health-IT-Innovation-in-Developing-Countries.aspx [Accessed 15 August 2009].

FrontlineSMS:Medic, 2009. *Products*. [Online] Available at: http://medic.frontlinesms.com/product-tour/ [Accessed 15 July 2010].

George, Dana, 2009. *Make Your Own Stethoscope*. How To. [Online] (Updated 26 February 2009) Available at: http://www.ehow.com/how_4810228_own-stethoscope.html [Accessed 10 December 2009].

Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE, 2000. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220; 2000 (June 13).

Graf, R., 1999. *Modern dictionary of electr*onics. 7th ed., Butterworth-Heinemann, Woburn, p. 161.

Greaves, I., Porter, K., and Burke, D., 1997. *Key topics in trauma.* BIOS Scientific Publishers Limited, Guildford, UK, p. 258.

Greer, I., 2001. *Mosby's color atlas and text of obstetrics and gynecology*. Mosby International Limited, China, p. 13.

GSMArena.com, 2007a. *GSMArena feature lab: Enter new goodies.* Reviews. [Online] (Updated 5 October 2007) Available at: http://www.gsmarena.com/latest_features-review-171p2.php [Accessed 7 July 2010].

GSMArena.com, 2007b. *Nokia N82 Review: Comes with xenon.* Reviews. [Online] (Updated 10 December 2007) Available at: http://www.gsmarena.com/nokia_n82-review-195p5.php [Accessed 3 March 2010].

GSMArena.com, 2008a. *HTC Touch HD Review: Windows almighty.* Reviews. [Online] (Updated 28 November 2008) Available at: http://www.gsmarena.com/htc_touch_hd-review-300p6.php [Accessed 7 July 2010].

GSMArena.com, 2008b. *LG KT610 review: Communicator wannabe.* Reviews. [Online] (Updated 27 August 2008) Available at: http://www.gsmarena.com/lg_kt610-review-276p5.php [Accessed 7 July 2010].

GSMArena.com, 2008c. *Nokia E66 review: Clean-cut business tool.* Reviews. [Online] (Updated 15 August 2008) Available at: http://www.gsmarena.com/nokia_e66-review-274p5.php [Accessed 7 July 2010].

GSMArena.com, 2009. *T-Mobile G1 Review: The whole cagoogle.* Reviews. [Online] (Updated 20 March 2009) Available at: http://www.gsmarena.com/t_mobile_g1-review-337p7.php [Accessed 3 March 2010].

GSMArena.com, 2010. *Motorola Milestone Review: Hello Droid.* Reviews. [Online] (Updated 24 January 2010) Available at: http://www.gsmarena.com/motorola_milestone-review-441p6.php [Accessed 3 March 2010].

Gut, U. and Bayerl, P., 2004. Measuring the reliability of manual annotations of speech corpora. *In Proceedings of the Second International Conference on Speech Prosody* (SP2004), March 23-26, Nara, Japan, pp. 565-568.

Hadjileontiadis, L. J., 2008. *Lung Sounds: An Advanced Signal Processing Perspective*, J. D. Enderle (Ed.), in Synthesis Lectures on Biomedical Engineering Series, San Francisco Bay Area, CA: Morgan & Claypool Publishers LLC, 2008, pp. 1-9.

Hadjileontiadis, L. J., Tolias, Y. A., and Panas, S. M., 2002, Intelligent system modeling of bioacoustic signals using advanced signal processing techniques, in: Intelligent Systems: Technologies and Applications – Vol.III: Signal, Image, and Speech Processing, C. T. Leondes, ed., CRC Press, Boca Raton, FL, chap. 3, pp. III 103-156.

Haggstrom, M., 2009. *File: Upper body front – Wikipedia, the free encyclopedia*. Wikimedia Commons [Online] (Updated 29 April 2009) Available at: http://en.wikipedia.org/wiki/File:Upper_body_front.png [Accessed 6 May 2010].

Hall-Clifford, R., 2009. *Oral Rehydration Therapy in Highland Guatemala: Long-Term Impacts of Public Health Intervention on the Self*, PhD Thesis, Boston University, Boston.

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.

Hansler, E. and Schmidt, G., 2008. *Speech and Audio Processing in Adverse Environments*, Series on Signals and Communication Technology, Springer, Berlin.

Harden, Scott, 2009. *DIY ECG Machine on the Cheap*. [Online] (Updated 14 August 2009) Available at: http://www.swharden.com/blog/category/diy-ecg-home-made-electrocardiogram/ [Accessed 13 January 2010].

Higgins, D., 2007. Reliability of human annotation of semantic roles in noisy text, *In Proceedings of ternational Conference on Semantic Computing*, Irvine, CA, pp. 501-508. (doi: 10.1109/ICSC.2007.102)

Hoffmann, F., Rosenbach, O., Aufrecht, E., and Stengel, A., 1902. *Diseases of the bronchi, lungs and pleura.* W.B. Saunders, p. 369.

Homaeinezhad, M., Atyabi, S., Deneshvar, E., Ghaffari, A., and Tahmasebi, M., 2010. Optimal Delineation of PCG Sounds via False-Alarm Bounded Segmentation of a Wavelet-Based Principal Components Analyzed Metric. *International Journal for Numerical Methods in Biomedical Engineering,* (in press).

HTC, 2009. *HTC – Products - T-mobile G1 Specification* [Online] Available at: http://www.htc.com/www/product/g1/specification.html [Accessed 9 July 2010].

Hughes, N., Tarassenko, L., and Roberts, S., 2003. Markov models for automated ECG interval analysis. *In Advances in Neural Information Processing Systems* 16, pp. 611–618, Cambridge, MA, 2003. MIT Press.

Hurtig-Wennlof, A., Ahlstrom, C., Egerlid, R., Resare, M., Ask, P., and Rask, P., 2009. Heart sounds are altered by open cardiac surgery. *Experimental and Clinical Cardiology*, 14(2), pp. 18-20.

iAuscultate, 2009. *iAuscultate.com* [Online] (Updated 2009) Available at: http://iauscultate.com/iAuscultate.html [Accessed 15 December 2009].

IBM and Microsoft, 1991. Multimedia Programming Interface and Data Specifications 1.0. IBM Corporation and Microsoft Corporation (Updated August 1991) [Online] Available at: http://www.tactilemedia.com/info/MCI_Control_Info.html

Independent Audio Benchmarking Project, 2009. *RightMark Audio Analyzer.* (Updated 23 March 2009) [Online] Available at: http://audio.rightmark.org/products/rmaa.shtml [Accessed 7 July 2010].

Istepanian, R. S. H., Laxminarayan, S. and Pattichis, C. S., 2006. Biosignals and Compression Standards in Topics in Biomedical Engineering: M-Health, Emerging Mobile Health Systems, Springer, US, pp 277-29. (doi: 10.1007/0-387-26559-7_21)

James, K., 2009. *Software Engineering.* PHI Learning Private Limited, New Delhi, p. 214.

Jardins, T., and Burton, G., 2005. *Clinical manifestations and assessment of respiratory disease.* Mosby Elsevier, St. Louis, p. 16.

Jimenez-Gonzalez, A. and James, C.J., 2009. Extracting sources from noisy abdominal phonograms: A single-channel blind source separation method. *Medical & Biological Engineering & Computing,* 47:6, pp. 655-664. (doi: 10.1007/s11517-009-0474-8)

Jin, Z., Oresko, J., Huang, S., and Cheng, A. C., 2009. HeartToGo: A Personalized medicine technology for cardiovascular disease prevention and detection. In *IEEE/NIH Life Science Systems and Applications Workshop (LiSSA 2009)*, pp. 80-83. (doi: 10.1109/LISSA.2009.4906714)

Joshi, N., 1999. The Third Heart Sound. *Southern Medical Journal.* 92(8), pp. 756-761.

Joshi, R. and Sanderson, A., 1999. *Multisensor fusion: a minimal representation framework.* World Scientific Publishing Co., Singapore.

Kaisla, T. S., Sovijarvi, A., Piirila, P., Rajala, H., Haltsonen, S., and Rosqvist, T., 1991. Validated method for automatic detection of lung sound crackles. *Medical and Biological Engineering and Computing Journal*, 29(5), pp. 517-521. (doi: 10.1007/BF02442324)

Kazama, S., 1990. A new stereophonic stethoscope. *Japanese Heart Journal,* **31,** pp. 837-843.

King, T. and Reiss, M., 2001. *Practical advanced biology*. Thomas Nelson and Sons Ltd, UK, p. 132.

Krajnik, M., Damps-Konstanska, I., Gorska, L., and Jassem, E., 2010. A portable automatic cough analyser in the ambulatory assessment of cough. *Biomedical Engineering Online*, 9:17. (doi: 10.1186/1475-925X-9-17)

Landois, L., 1891. *A Text-book of human physiology*. 4th ed., Vol 1, Charles Griffin and Company, London, pp. 206-208.

Lewis, C., 1999. Every Breath You Take: Preventing and Treating Emphysema. *FDA Consumer Magazine,* U.S. Food and Drug Administration, (March-April).

Loudon, R. and Murphy, R., 1984. *State of the art: Lung sounds.* Am Rev Respir Dis, 130, pp. 663-673.

M-Audio, 2010. *Fast Track Pro – 4x4 Mobile USB Audio/MIDI Interface with Preamps*. [Online] Available at: http://www.m-audio.com/products/en_us/FastTrackPro.html [Accessed 7 July 2010].

Mackay, J., Mensah, G., Mendis, S., and Greenlund, K., 2004. *The atlas of heart disease and stroke.* World Health Organization, Geneva.

Malarvili, M., Kamarulafizam, I., Hussain, S., and Helmi, D., 2003. Heart Sound Segmentation Algorithm Based on Instantaneous Energy of Electrocardiogram. *Computers in Cardiology Journal*, 30, pp. 327-330.

Malkin, R. A., 2007. Technologies for clinically relevant physiological measurements in developing countries. *Physiol. Meas.,* 28**,** pp. 57-63. (doi: 10.1088/0967-3334/28/8/R01)

Martínez, J.P., Olmos, S., and Laguna, P., 2000. Evaluation of a wavelet-based ECG waveform detector on the QT Database. *XXVII International Conference on Computers in Cardiology, IEEE Computer Society*, Boston, pp. 81-84.

Mathew, K., and Aggarwal, P., 2008. *Medicine: Prep Manual for Undergraduates*. 3rd edn, Reed Elsevier India Private Limited, New Delhi, p. 410.

MathWorks 2009, *MATLAB*, ver. 7.8.0.347, computer program, The MathWorks Inc., Natick, MA, USA.

Mazeika, G. G., and Swanson, R., 2007. *Respiratory Inductance Plethysmography: An Introduction* [Report] Mukilteo, WA: Pro-Tech Services, Inc.

McCann, J., Holmes, H., Robinson, J., Putterman, A., et al., 2002. *Illustrated manual of nursing practice.* 3rd ed., Lippincott Williams & Wilkins, U.S.

McCann, J., Holmes, H., Robinson, J., Putterman, A., et al., 2005a. *Assessment: a 2-in-1 reference for nurses.* Lippincott Williams & Wilkins, Springhouse, Pennsylvania, p. 109.

McCann, J., Moreau, D., Robinson, J., Putterman, A., et al., 2005b. *Breath Sounds Made Incredibly Easy!* Lippincott Williams & Wilkins, Pennsylvania.

McCann, J., Moreau, D., Robinson, J., Putterman, A., et al., 2005c. *Heart Sounds Made Incredibly Easy.* Lippincott Williams & Wilkins, U.S.

Meditec, Inc., 2009. *Heart Examination: Heart Sounds and Review of Fundamentals* [Online] Available at: http://www.meditec.com/resourcestools/medical-reference-links/heart-examination/ [Accessed 13 December 2009].

MedlinePlus, 2009. *Idiopathic pulmonary fibrosis: MedlinePlus Medical Encyclopedia* [Online] Service of the U.S. National Library of Medicine and the National Institutes of

Health, Available at: http://www.nlm.nih.gov/medlineplus/ency/article/000069.htm [Accessed 13 June 2010].

Menchi, D., 2008. Pesinet – Telediagnosis for children. Lyon Conference for Digital Solidarity. [Online] Available at: http://www.digitalsolidarity2008.org/en/The-conference/News/Pesinet-Telediagnosis-for-children [Accessed 20 October 2008].

Microsoft, 1994. Microsoft Multimedia Standards Update: New Multimedia Data Types and Data Techniques (Updated 15 April 1994) [Online] Available at: http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/ WAVE/Docs/RIFFNEW.pdf

MIT OpenCourseWare, 2003a. 6.857 Network and Computer Security. Electrical Engineering and Computer Science. [Online] Available at: http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-857-network-and-computer-security-fall-2003/ [Accessed 2 July 2010].

MIT OpenCourseWare, 2003b. HST.950J / 6.872J Medical Computing. Health Sciences and Technology. [Online] Available at: http://ocw.mit.edu/courses/health-sciences-and-technology/hst-950j-medical-computing-spring-2003/ [Accessed 2 July 2010].

MIT OpenCourseWare, 2008. MAS.965 / 6.976 / SP.716 NextLab I: Designing Mobile Technologies for the Next Billion Users. Media Arts and Sciences Courses. [Online] Available at: http://ocw.mit.edu/courses/media-arts-and-sciences/mas-965-nextlab-i-designing-mobile-technologies-for-the-next-billion-users-fall-2008/ [Accessed 2 July 2010].

Moody, G. B., Koch, H. and Steinhoff, U., 2006. The PhysioNet/Computers in Cardiology Challenge 2006: QT Interval Measurement, *Computers in Cardiology,* 33, pp 313-316.

Moody, G.B. and Lehman, L.H., 2009. Predicting acute hypotensive episodes: The 10th annual PhysioNet/Computers in Cardiology Challenge *Computers in Cardiology*, 36 pp 541-544.

Moussavi, Z., 2006. *Fundamentals of Respiratory Sounds and Analysis,* J. D. Enderle (Ed.), in Synthesis Lectures on Biomedical Engineering Series, Manitoba, Canada: Morgan & Claypool Publishers LLC, pp. 1-34.

Murphy, R., Del Bono, E., Davidson, F., 1989. Validation of an automatic crackle (rale) counter. *Am Rev Respir Dis,* 140(4), pp. 1017-1020.

Murray, J., Nadel, J., 1995. *Textbook of Respiratory Medicine*. Saunders, Philadelphia, pp. 563-584.

Nemati, S., Malhotra, A., and Clifford G. D., 2010. Data Fusion for Improved Respiration Rate Estimation*, EURASIP Journal on Advances in Signal Processing 2010*, Article ID 926305, 8 May. (doi: 10.1155/2010/926305)

Neamatullah, I., Douglass, M. M., Lehman, L. H., Reisner, A., Villarroel, M., Long, W. J., Szolovits, P., Moody, G. B., Mark, R. G. and Clifford, G. D., 2008. *Automated De-Identification of Free-Text Medical Records,* BMC Medical Informatics and Decision Making 2008, 8:32 (24 July). (doi: 10.1186/1472-6947-8-32)

Nurse, D. *Philippines Health System* [Report] Manila: Healthcare Information and Management Systems Society (HIMSS), EHR Global Task Force, 2009.

OpenMRS, 2007. *OpenMRS data model version 1.10*. (Updated 20 April 2007) [Online] Available at: http://openmrs.org/images/5/59/Openmrs_data_model_1.10.png?format=raw [Accessed 30 July 2009].

OpenMRS, 2008. *Concept Dictionary*. (Updated 30 July 2008) [Online] Available at: http://openmrs.org/wiki/Concept_dictionary [Accessed 20 July 2009].

OpenMRS, 2009a. *Google Maps Image Viewer Module for OpenMRS*. (Updated 19 May 2009) [Online] Available at: http://openmrs.org/wiki/Google_Maps_Image_Viewer_Module [Accessed 19 June 2009].

OpenMRS, 2009b. *Media Viewer Module for OpenMRS*. (Updated 10 June 2010) [Online] Available at: http://openmrs.org/wiki/Google_Maps_Image_Viewer_Module [Accessed 29 June 2010].

OpenMRS, 2010a. *OpenMRS Medical Record System*. (Updated 8 July 2010) [Online] Available at: http://openmrs.org/wiki/OpenMRS [Accessed 10 October 2009].

OpenMRS, 2010b. Usage Statistics Module for OpenMRS. (Updated 6 May 2010) [Online] Available at: http://openmrs.org/wiki/Usage_Statistics_Module [Accessed 7 May 2010].

Osborn, C., 2000. *Statistical applications for health information management.* Aspen Publishers, Maryland.

Pan, J. and Tompkins, W.J, 1985. A real-time QRS detection algorithm. IEEE Trans. *Biomed. Eng.,* BME, 32 (3), pp. 230-236. (doi: 10.1109/TBME.1985.325532)

Pandey, M. et al., 1989. Air Pollution in Developing Countries and Acute Respiratory Infection in Children. *The Lancet,* (25 February).

Pasterkamp, H., et al., 2008. *R.A.L.E. Lung Sounds 3.12 Professional Edition* [Software] Winnipeg, Canada: PixSoft Inc.

Pasterkamp H., Carson C., Daien D., Oh Y., 1989. Digital respirosonography. New images of lung sounds. *Chest*, 96(6), pp. 1405-12 (doi: 10.1378/chest.96.6.1405)

Philips Medical Systems, 2009. *Fighting Pneumonia*. Philanthropy by Design 2009. [Online] Available at: http://www.design.philips.com/philips/shared/assets/design_assets/pdf/nvbD/august2009/Fighting_pneumonia.pdf [Accessed 1 October 2009].

Popular Science, 1959. What you should know about sheet polyethylene – The plastic film of countless uses. Materials File, 175: 5, Bonnier Corporation, pp. 163-165.

Rao, C., and Guha, S., 2001. *Principles of medical electronics and biomedical instrumentation*, Hyderabad, Universities Press (India) Limited, p. 145.

Richard, W., Zar, D., and Solek, R., 2008. A Low-Cost B-Mode USB Ultrasound Probe, *Ultrasonic Imaging*, 30, pp. 21-28.

RnCeus Interactive LLC, 2006. *Normal Breath Sounds* [Online] Available at: http://www.rnceus.com/resp/respnorm.html [Accessed 13 November 2009].

Rosdahl, C. and Kowalski, M., 2008. *Textbook of basic nursing.* pp. 521-522, 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins.

Scanlon, V. C. and Sanders T., 1995. *Essentials of Anatomy and Physiology*. 4th ed. Philadelphia, PA: F.A. Davis Co.

Schmidt, S., Holst-Hansen, C., Graff, C., Toft, E., and Struijk, J., 2010. Segmentation of heart sound recordings by a duration-dependent hidden Markov model. *Physiological Measurement*, 31:4, pp. 513-529. (doi: 10.1088/0967-3334/31/4/004)

Science Fair Adventure, 2007. *Home-made Stethoscope*. Science Project Details. [Online] Available at: http://www.sciencefairadventure.com/ProjectDetail.aspx?ProjectID=165 [Accessed 5 December 2009].

Seidel, H.M., 1995. *Mosby's Guide to Physical Examination*. 3rd ed. St. Louis, U.S.: Yearbook Inc, p. 69.

Sheldon, P., and Doe, J.. *The Development of the Stethoscope: An Exhibition Showing the Work of Lannec and His Successors Held in the Library February 13 to March 29, 1935*. Bull NY Acad Med 1935: NCBI, pp. 608-26.

Sherwood, L., 2009. *Human Physiology: From Cells to Systems*. Brooks/Cole, Cengage Learning, Belmont, p. 321-323.

Smeltzer, S., Bare, B., Hinkle, J., and Cheever, K., 2009. *Brunner and Suddarth's textbook of medical-surgical nursing.* Unit 7 Digestive and Gastrointestinal Function, Lippincott Williams & Wilkins, p. 574.

Smith, S., 1997. *The scientist and engineer's guide to digital signal processing.* Chapter 16 Windowed-Sinc Functions, California Technical Pub.

Souhami, R. and Moxham, J., 2002. *Textbook of medicine*. 4th ed., Elsevier Health Sciences, p. 622.

Spector, J., Daga, S., 2008. Preventing those so-called stillbirths. *Bulletin of the World Health Organization*, 86:4 (April), pp. 241-320.

Spyropoulos, B., Tzavaras, A., Botsivaly, M., and Koutsourakis, K., 2008. Supporting Cardio-Respiratory Diseases Related House-Call Medical Practice. In *CeHR Conference Proceedings 2007*, Berlin, p. 157-168.

Stethographics, Inc., 2007. *Product Downloads: Sound Samples* [Online] Physiology, Available at: http://www.stethographics.com/main/products_downloads.html [Accessed 10 June 2010].

Stethoscope Basics – What makes a Good Stethoscope. Reviews & Guides. [Online] (Updated 2 April 2010) Available at: http://reviews.ebay.com/Stethoscope-Basics-What-makes-a-Good-Stethoscope_W0QQugidZ10000000004429816 [Accessed 5 May 2010].

Syed, Z., 2003. *MIT Automated Auscultation System*, MEng Thesis, Massachusetts Institute of Technology, Cambridge. (uri: http://hdl.handle.net/1721.1/18018)

Tan, A., 2009. *Smartphone-Based Fetal Monitors Could Save Lives in Remote Areas*. Microsoft Research. [Online] Available at: http://research.microsoft.com/enus/collaboration/focus/health/smartphone_fetal_monitor.aspx [Accessed 20 September 2009].

Tarassenko, L., Townsend, N., Clifford, G., Mason, L., Burton J., and Price, J., 2001. Medical signal processing using the software monitor. In *DERA/IEE Workshop Intelligent Sensor Processing*, London, UK, no. 3 in 140, pp. 1–4.

Taylor, C., 2000. *The Kingfisher Science Encyclopedia*. Hong Kong: Kingfisher Publications Plc.

Trimline Medical Products, 2004. *Stethoscope Basics: Parts of a Stethoscope*. [Brochure] Available at: http://www.trimlinemed.com/downloads/pdfs/stethoscopebasics.pdf [Accessed 5 December 2009].

Tsui, J., 2004. *Digital Techniques for Wideband Receivers.* 2nd edn, SciTech Publishing, Raleigh, p. 262.

UNICEF, 2006. *Pneumonia: The Forgotten Killer of Children* [Report]. World Health Organization and United Nations Children's Fund. Available at: http://www.unicef.org/publications/files/Pneumonia_The_Forgotten_Killer_of_Children.pdf [Accessed 14 June 2010].

UNICEF, 2008. *Acute Respiratory Infection* [Report D7341] Special Session of the UN General Assembly on Children, 8-10 May 2008. Available at: http://www.unicef.org/specialsession/about/sgreport-pdf/08_AcuteRespiratoryInfection_D7341 Insert_English.pdf [Accessed 14 November 2009].

Unified Medical Language System (UMLS), 2008. *UMLS Enhanced VA/KP Problem List Subset of SNOMED CT.* United States National Library of Medicine. National Institutes of Health. (Updated 13 July 2008) [Online] Available at: http://www.nlm.nih.gov/research/umls/Snomed/snomed_problem_list.html

Unified Medical Language System (UMLS), 2010. *UMLS SNOMED Clinical Terms (SNOMED CT).* United States National Library of Medicine. National Institutes of Health. (Updated 12 May 2010) [Online] Available at: http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html

United Nations, 2009. *Improve Maternal Health*. United Nations Millenium Development Goals [Online] Available at: http://www.un.org/millenniumgoals/maternal.shtml [Accessed 15 August 2009]

University of Maryland Medical Center, 2006. *Respiratory Disease: Interstitial Lung Diseases (Pulmonary Fibrosis)* [Online] (Updated 8 June 2006) Available at: http://www.umm.edu/respiratory/interstit.htm [Accessed 14 June 2010].

University of Virginia Health System, 2008. *Vital Signs (Body Temperature, Pulse Rate, Respiration Rate, Blood Pressure.* Non-Traumatic Emergencies. [Online] (Updated 24 January 2008) Available at: http://www.healthsystem.virginia.edu/uvahealth/adult_nontrauma/vital.cfm [Accessed 20 November 2009].

University of Washington Department of Medicine. *Technique: Heart Sounds and Murmurs*. Advanced Physical Diagnosis Learning and Teaching at the Bedside, Skill Modules, Techniques. [Online] Available at: http://depts.washington.edu/physdx/heart/tech.html [Accessed 20 November 2009].

UrbanSTEW, 2009. *Rehearsal Assistant Platform.* [Online] Available at: http://urbanstew.org/_rehearsalassistant/#Rehearsal_Assistant_platform [Accessed 20 April 2010].

U.S. Department of Health and Human Services, 2003. *The Security Rule*. Health Information Privacy (Updated 20 February 2003) [Online] Available at: http://www.hhs.gov/ocr/privacy/hipaa/ administrative/securityrule/index.html [Accessed 30 April 2010].

Utts, J., Heckard, R., 2006. *Statistical Ideas and Methods.* Thomson Brooks/Cole, Belmont, p. 619.

Varga, A., Steenneken, J., Tomlinson, M., and Jones, D., 1992. *The NOISEX-92 study on the effect of additive noise on automatic speech recognition*, Tech. Rep. DRA Speech Res. Unit. Malvern, Worcestershire, U.K.

Vital Wave Consulting and United Nations Foundation, 2009. mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World. [Report] UN Foundation-Vodafone Foundation Partnership. Washington, D.C. and Berkshire, UK.

Walter, A., Rutledge, M., Edgar, C., and Davis, R., 2004. *First Responder Handbook: Law enforcement edition.* Thomson Delmar Learning, Clifton Park, p. 164.

Wang, F., Syeda-Mahmood, T., Beymer, D., 2007. Finding Disease Similarity by Combining ECG with Heart Auscultation Sound. *Computers in Cardiology*, 34, pp. 261-264.

Weil, S., 2004. *HIPAA Security Rule*. Symantec Connect (Updated 2 March 2004) [Online] Available at: http://www.symantec.com/connect/articles/hipaa-security-rule [Accessed 30 April 2010].

White, G., 2003. *Basic Clinical Lab Competencies for Respiratory Care: An Integrated Approach*, 4th ed., Thomson Delmar Learning, New York, pp. 19 – 20.

White, L., 2004. *Foundations of Nursing*. 2nd ed., Thomson Delmar Learning, pp. 848-892.

Whitehead, M., Dahlgren, G., and Evans, T., 2001. Equity and health sector reforms: can low-income countries escape the medical poverty trap? *The Lancet*, 358, pp. 833-836.

Win Health, 2010. *Electronic Stethoscopes* [Online] Brockhirst, Oxnam Road, Jedburgh. Available at: http://www.win-health.com/electronic-digital-stethoscopes.html [Accessed 3 April 2010].

Worden K. and Tomlinson, G., 2001. *Nonlinearity in structural dynamics: detection, identification, and modeling*, Institute of Physics Publishing, UK, p. 127.

World Health Organization (WHO), 2007. *International medical guide for ships: including the ship's medicine chest.* 3rd ed., WHO Press, Geneva, pp. 139-149.

World Health Organization (WHO), 2008a. *Asthma.* Fact Sheet Nº307. [Online] (Updated May 2008) Available at: http://www.who.int/mediacentre/factsheets/fs307/en/index.html

World Health Organization (WHO), 2008b. *World Health Statistics 2008.* WHO Press, France.

World Health Organization, 2009a. *Acute Respiratory Infections.* Initiative for Vaccine Research. [Online] (Updated September 2009) Available at: http://www.who.int/vaccine_research/diseases/ari/en/index.html

World Health Organization, 2009b. *Cardiovascular diseases (CVDs).* Fact Sheet Nº317. [Online] (Updated September 2009) Available at: http://www.who.int/mediacentre/factsheets/fs317/en/index.html

World Health Organization, 2009c. *Chronic obstructive pulmonary disease (COPD).* Fact Sheet Nº315. [Online] (Updated November 2009) Available at: http://www.who.int/mediacentre/factsheets/fs315/en/index.html

World Health Organization, 2010. *Chronic respiratory diseases.* Programmes and projects. [Online] Available at: http://www.who.int/respiratory/en/

World Health Organization Africa, 2009. *Chronic respiratory diseases prevention and control.* Regional Office for Africa, Non Communicable Disease Prevention & Control (NPC) Programme components. [Online] Available at: http://www.afro.who.int/en/divisions-a-programmes/ddc/non-communicable-diseases-managementndm/programme-components/chronic-respiratory-diseases.html

Woyke, E., 2009. Nokia's Gigantic App Store [Online] (Updated 7 May 2009) Forbes Available at: http://www.forbes.com/2009/05/07/nokia-ovi-store-technology-wireless-nokia.html [Accessed 8 July 2010].

Ye, N., 2004. *The Handbook of Data Mining.* 2nd edn, Lawrence Erlbaum Associates, New Jersey, p. 283.

Yi, G., 2004. *A Software Toolkit for Acoustic Respiratory Analysis,* MEng Thesis, Massachusetts Institute of Technology, Cambridge. (uri: http://hdl.handle.net/1721.1/33391)

Yip, L., and Zhang, Y., 2001. *Reduction of heart sounds from lung sound recordings by automated gain control and adaptive filtering techniques*. [Report] Hong Kong: Chinese University of Hong Kong Shatin Department of Electrical Engineering. (uri: http://handle.dtic.mil/100.2/ADA410078)