THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Automated Methods for Cough Assessments and Applications in Screening Paediatric Respiratory Diseases

Yusuf Aziz Amrulloh

BEng, MEng

A thesis submitted for the degree of Doctor of Philosophy at The University of Queensland in 2014

School of Information Technology and Electrical Engineering

Abstract

Cough has the potential to be developed as a diagnostic tool; however, it has not yet been fully explored. Researchers have attempted to study cough in adults, but none of these studies involved paediatric populations younger than five years of age with respiratory diseases such as pneumonia. The study of automated wet/dry cough classifications in paediatric populations is not yet available either. At the moment, the identification of wet/dry cough is carried out manually by physicians. The results of this process are subjective and depend on the skills and experiences of the observers. Cough is one of the main symptoms of pneumonia, but the World Health Organization (WHO) only uses its existence for screening in pneumonia. An acoustic analysis of the pneumonia cough sounds for diagnosing this disease has not yet been explored. Further, quantitative study of cough analysis is still immature; physicians still have to identify and listen to the cough manually, which is a tedious and time-consuming task. An automated method capable of segmenting cough sound from recordings is urgently required.

This thesis proposes the development of innovative cough sound analysis based methods to address the problem of wet/dry cough classification, substituting the bronchodilator test in resource-limited settings and segmenting cough from recordings automatically. In my approach, the cough samples were collected using non-contact sensors at a hospital in a developing country. All subjects included in this thesis are members of the paediatric population suffering from respiratory diseases such as pneumonia and asthma.

The supports for my work are the results from preliminary studies and the pathophysiology of respiratory diseases. The infections stimulate the excessive production of mucus in the airways. In pneumonia, the mucus also fills the alveoli and causes lung consolidation. The opening and closing of the collapsed alveoli/airways produce crackle sounds. I hypothesize that the vibration of mucus, the inflammation of airways, the lung consolidation, and the crackle sounds alter the acoustic of pneumonia cough sounds such that they are distinguishable from the coughs of other diseases such as asthma.

To capture the cough sound signatures, I extracted features such as non-Gaussianity score, Mel-frequency cepstral coefficients, Shannon entropy, formant frequency, and zero crossing rates. These features were used for training classifiers to classify wet/dry coughs automatically, to differentiate pneumonia from asthma, and to segment cough from recordings. My results show that the proposed methods achieve high performance for the designed purposes.

This thesis contributes to the development of a pioneering class of technology that addresses fundamental gaps in cough sound analysis. The non-contact technology for cough analysis is perfectly matched for children. In addition, it does not require elaborate sterilization process efforts. The automated wet/dry cough classification method facilitates objective cough assessment, and is useful for long-term wet/dry cough study. My cough based method for separating pneumonia and asthma can revolutionize the diagnosis of these diseases in limited-resource settings. The method can be developed into an affordable system for replacing the bronchodilator test. Further, my automated segmentation method has the potential to be developed as a cough counting device, as well as the front end of cough analysis systems.

For future work, my methods should be tested in a larger dataset to develop a robust system. The methods can be developed for smart phone application or deployed in a low cost embedded system.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the General Award Rules of The University of Queensland, immediately made available for research and study in accordance with the *Copyright Act 1968*.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

1. Y. A. Amrulloh, U. R. Abeyratne, V. Swarnkar, D. Herath, R. Triasih, and A. Setyati, “Separating asthma from pneumonia in resource limited areas: Can cough sound analysis replace the bronchodilator test for children”, Physiological Measurement, UK, 2014, (under review).
2. Y. A. Amrulloh, U. R. Abeyratne, V. Swarnkar, R. Triasih, A. Setyati, “Automatic cough segmentation from non-contact sound recordings in paediatric wards”, Biomedical Signal Processing and Control, Netherland, 2014 (submitted).
3. V. Swarnkar, U. R. Abeyratne, Y. Amrulloh, “Neural network based algorithm for automatic identification of cough sounds”, 35th Annual International Conference of the IEEE EMBS, Osaka, Japan, July 2013.
4. V. Swarnkar, U. R. Abeyratne, A. B. Chang, Y. A. Amrulloh, R. Triasih, A. Setyati, and R. Triasih, “Automatic identification of wet and dry cough in paediatric patients with respiratory diseases”, Annals of Biomedical Engineering, New York: Springer, DOI: 10.1007/s10439-013-0741-6, January 2013.
5. V. Swarnkar, U. R. Abeyratne, Y. A. Amrulloh, A. B. Chang, “Automated algorithm for wet/dry cough sounds classification”, 34th Annual International Conference of the IEEE EMBS, San Diego, California, 28 August – 1 September 2012, pp. 3147-3150.

Publications included in this thesis

1. Y. A. Amrulloh, U. R. Abeyratne, V. Swarnkar, D. Herath, R. Triasih, A. Setyati, "Separating asthma from pneumonia in resource limited areas: Can cough sound analysis replace the bronchodilator test for children?", *Physiological Measurement*, UK, 2014, (under review).

Contributors	Statement of contribution	
Yusuf A. Amrulloh (Candidate)	Concept and design	(80%)
	Interpretation of data	(80%)
	Computations	(80%)
	Drafting and writing	(60%)
Udantha Abeyratne	Concept and design	(20%)
	Interpretation of data	(20%)
	Drafting and writing	(30%)
Dulip Herath	Computations	(20%)
Vinayak Swarnkar	Drafting and writing	(10%)
Rina Triasih	Clinical data interpretation	(50%)
Amalia Setyati	Clinical data interpretation	(50%)

6. Y. A. Amrulloh, U. R. Abeyratne, V. Swarnkar, R. Triasih, A. Setyati, "Automatic cough segmentation from non-contact sound recordings in paediatric wards", *Biomedical Signal Processing and Control*, Netherland, 2014 (submitted).

Contributors	Statement of contribution	
Yusuf A. Amrulloh (Candidate)	Concept and design	(80%)
	Interpretation of data	(80%)
	Computations	(100%)
	Drafting and writing	(60%)
Udantha Abeyratne	Concept and design	(20%)
	Interpretation of data	(20%)
	Drafting and writing	(30%)
Vinayak Swarnkar	Drafting and writing	(10%)
Rina Triasih	Clinical data interpretation	(50%)

Amalia Setyati	Clinical data interpretation	(50%)

2. V. Swarnkar, U. R. Abeyratne, Y. A. Amrulloh, "Neural Network Based Algorithm for Automatic Identification of Cough Sounds", *35th Annual International Conference of the IEEE EMBS*, Osaka, Japan, July 2013.

Contributors	Statement of contribution	
Vinayak Swarnkar	Concept and design	(40%)
	Interpretation of data	(40%)
	Computations	(50%)
	Drafting and writing	(40%)
Udantha Abeyratne	Concept and design	(20%)
	Interpretation of data	(20%)
	Computations	(10%)
	Drafting and writing	(30%)
Yusuf A. Amrulloh (Candidate)	Concept and design	(40%)
	Interpretation of data	(40%)
	Computations	(40%)
	Drafting and writing	(30%)

3. V. Swarnkar, U. R. Abeyratne, A. B. Chang, Y. A. Amrulloh, R. Triasih, A. Setyati, "Automatic identification of wet and dry cough in pediatric patients with respiratory diseases", *Annals of Biomedical Engineering*, New York: Springer, DOI: 10.1007/s10439-013-0741-6, January 2013.

Contributors	Statement of contribution	
Vinayak Swarnkar	Concept and design	(40%)
	Interpretation of data	(40%)
	Computations	(50%)
	Drafting and writing	(40%)
Udantha Abeyratne	Concept and design	(20%)
	Interpretation of data	(20%)
	Computations	(10%)
	Drafting and writing	(30%)

Contributors	Statement of contribution	
Yusuf A. Amrulloh (Candidate)	Concept and design Interpretation of data Computations Drafting and writing	(40%) (40%) (40%) (30%)
Anne B. Chang	Scoring wet/dry cough sounds	(33.33%)
Rina Triasih	Scoring wet/dry cough sounds	(33.33%)
Amalia Setyati	Scoring wet/dry cough sounds	(33.33%)

4. V. Swarnkar, U. R. Abeyratne, Y. A. Amrulloh, A. B. Chang, "Automated algorithm for wet/dry cough sounds classification", *34th Annual International Conference of the IEEE EMBS*, San Diego, California, 28 August – 1 September 2012, pp. 3147 – 3150.

Contributors	Statement of contribution	
Vinayak Swarnkar	Concept and design Interpretation of data Computations Drafting and writing	(40%) (40%) (50%) (50%)
Udantha Abeyratne	Concept and design Interpretation of data Drafting and writing	(20%) (20%) (30%)
Yusuf A. Amrulloh (Candidate)	Concept and design Interpretation of data Computations Drafting and writing	(40%) (40%) (50%) (20%)
Anne B. Chang	Scoring wet/dry cough sounds	(100%)

Contributions by others to the thesis

None

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgments

I would like to acknowledge The University of Queensland (UQ) for its financial and academic supports, particularly for the awards of the University of Queensland International Scholarships (UQI), the University of Queensland Research Scholarship (UQRS), and the UQ Indonesian Partnership (UQIPA), which made it possible to complete my PhD program.

I wish to express my heartfelt appreciation to my principal supervisor, Associate Professor Udantha Abeyratne, for sharing his wealth of knowledge with me, for providing a constructive research atmosphere, and for his endless support at all stages of my study. Also, I would like to express my deep gratitude and respect to him for being a person open to ideas, and for encouraging me and helping me to shape my own interests and ideas.

My sincere thanks go out to my co-supervisor Dr. Vinayak Swarnkar, whose advice and insight were extremely valuable for this thesis. I am grateful for all I learned from him and for the feedback he provided to improve my work.

I wish to express my sincere gratitude to my colleagues from Gadjah Mada University, Dr. Rina Triasih, Dr. Amalia Setyati, and Dr. Roni Naning, for giving me the opportunity to do research in the Pediatric Department at Sardjito Hospital, Yogyakarta, Indonesia. Their medical expertise and input have been invaluable in the construction of this thesis. I would also like to thank to Dr. Vicka Oktaria, Dr. Dwi Setyoreni, Septi Ayu, Hasfran, Pressi and all the staff of paediatric ward at Sardjito Hospital for their assistance with data acquisition. This thesis would not have been possible without their helps and support.

I would like to thank to my colleagues in the Sleep and Respiratory Systems Engineering group, Dulip Herath, Shaminda de Silva, Keegan Kosasih, and Akhter Shahin, for the many fruitful discussions, varied opinions, and constructive criticisms. I extend my sincere thanks to Ranveer Matharu for his unquantifiable assistance.

I would like to express my great appreciation to my parents for their continuous love and support. Above all, I would like to thank to my wife and daughters for their unwavering support and patience. I am indebted to them for all their sacrifices and hardships during the completion of my PhD degree. Your spirits lifted mine and gave me the strength to complete my study.

Keywords

Cough sound analysis, cough characteristics, paediatric pneumonia, bronchodilator test, cough segmentation

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC: code	080109	Pattern Recognition and Data Mining	: 45%
ANZSRC: code	090399	Biomedical Engineering not elsewhere classified	: 10%
ANZSRC: code	090609	Signal Processing	: 45%

Fields of Research (FoR) Classification

FoR: code	0903	Biomedical Engineering	: 30%
FoR: code	0906	Electrical and Electronic Engineering	: 70%

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANN	Artificial Neural Network
BGS	Bispectrum Score
COPD	Chronic obstructive pulmonary diseases
CG	Cough
CRP	C-reactive protein
dB	Decibel
EMG	Electromyography
ERS	European Respiratory Journal
FF	Formant frequencies
GMM	Gaussian Mixture Model
GERD	Gastroesophageal reflux
HPF	High Pass Filter
HMM	Hidden Markov Model
HACC	Hull Automatic Cough Counter
IMCI	Integrated Management of Childhood Illnesses
Kurt	Kurtosis
LCM	Leicester cough monitor
LOV	Leave one out cross validation
LogE	Log energy
LRM	Logistic-regression model
LRTI	Lower respiratory tract infection
MFCCs	Mel-frequency cepstral coefficients
MDD	Model design dataset
NPV	Negative predicted value
NGS	Non-Gaussianity score
NC	Non-cough
P	Pitch
PDF	Probability density function
PSS	Power spectral subtraction

PCI	Pneumonic Cough Index
PPV	Positive predicted value
PSD	Prospective study dataset
ROC	Receiver-operating curve
RSV	Respiratory syncytial virus
SH	Shannon entropy
SNR	Signal to noise Ratio
TDNN	Time delay neural network
UNICEF	United Nation International Children's Funds
URTI	Upper respiratory tract infection
WHO	World Health Organization
ZCR	Zero crossing rates

List of Symbols

f_c	cut off frequency
f_s	sampling rate
$s[n]$	discretised recording sound signals
$\hat{s}[n]$	filtered recording sound signals
$s_y[n]$	recorded audio sound signals
$\hat{s}_y[n]$	filtered audio sound signal
$s_{cg}[n]$	cough sound signals/cough events
$\hat{s}_{cg}[n]$	filtered cough sound signals /cough events
$s_{nc}[n]$	non-cough sound signals
$\hat{s}_{nc}[n]$	filtered non-cough sound signal
$\hat{s}_k[n]$	k^{th} sub-block/sub-segment having length N from signal $\hat{s}[n]$
N	length of rectangular window
k	sub-block/sub-segment
Φ	Mel-frequency cepstral coefficients
$\Delta\Phi$	first order differential of MFCC
$\Delta\Delta\Phi$	second order differential of MFCC
w	Hamming window
ζ_k	Fourier transform of kth sub-block
$S[m]$	Mel-filter banks
m	number of Mel filter banks
H_m	transfer function of Mel-filter
f_l	lowest frequencies of the Mel-filter bank in Hz
f_h	highest frequencies of the Mel-filter bank in Hz
$f[m]$	boundary points of the Mel-filter banks
B	Mel-scale
B^{-1}	inverse of Mel-scale
η	number of cepstral coefficients
γ	inverse of normal cumulative distribution function
ϑ	probability function
q_1	first quartile

q_3	third quartile
ψ	non-Gaussianity score
Z	zero crossing rates
$\hbar$	Shannon entropy
E	log energy
ε	arbitrarily small positive constant
λ_k	kurtosis
C	cumulant
B	bispectrum
P	one-dimensional slice of bispectrum
ξ	bispectrum score
κ	Kohen's Kappa
Ŋ	number of subjects
Ŋ1	number of subjects in model design dataset
Ŋ2	number of subjects in prospective dataset
Ç	number of cough events/episodes
Ç1	total number of cough events in model design dataset
Ç2	total number of cough events in prospective study dataset
Ç11	total number of cough events in model design dataset agreed by two scorers
MDD1	sub-set of model design dataset with Ç11 cough events
M_{MDD1}	feature matrix of cough events from dataset MDD1
M'_{MDD1}	selected feature matrix of cough events from dataset MDD1
M_{PSD}	feature matrix of cough events from dataset PSD having size Ç2 × F
M'_{PSD}	selected feature matrix of cough events from dataset
F	feature vectors
F_S	sub-set of selected features from F
Y	dependent output variable from a LRM model
$Y_{\Re}$	dependent output variable from model ℜ
θ	optimum decision threshold for Y
f_F	element of feature vectors
β_0	intercept in LRM
β_b	regression coefficient of independent variables in LRM
$Ł_{Ç11}$	number of LRM models developed from dataset Ç11
κ	Cohen's Kappa

p	significant value of a feature in the development of model
p_{ths}	selected p threshold
$\Re$	selected LRM model
$\theta_{\Re}$	selected decision threshold for $\Re$
μ	mean
σ	standard deviation
$\hat{c}[n]$	cough signal after HPF, PSS and pre-emphasis filters
Γ	MFCC feature matrix of a cough episode $\hat{c}_k[n]$
i	number of states in HMM model
j	number of Gaussian mixture model
$b_i(\Phi_k)$	probability density function of Φ_k from state i.
m_{ij}	j^{th} mixture weight of the i^{th} state of HMM
μ_{ij}	mean matrix of the j^{th} Gaussian distribution of the i^{th} state
Σ_{ij}	covariance matrix of the j^{th} Gaussian distribution of the i^{th} state
∂	dimension of the feature vectors
Λ	set of respiratory diseases classes of the cough episodes
λ_1	HMM model for pneumonia
λ_2	HMM model for asthma
LL	log-likelihood of a HMM model
$\hat{\varsigma}$	best pneumonia/asthma model that give the highest log-likelihood
Ω	a set of cough episodes from a patient
G_{MDD}	feature matrix from MDD
G_{PSD}	feature matrix from PSD
$u_m[k]$	output of neural network (TDNN)
$\tilde{u}_m[k]$	output of moving average filter
β	tap length of a moving average filter
ρ	threshold for the output of moving average filter
q	number of consecutive sub-blocks
τ_q	total time duration of q consecutive sub-blocks with $\gamma = 1$
τ_{min}	minimum cough sound durations computed from the cough events in MDD
τ_{max}	maximum cough sound durations computed from the cough events in MDD
δ	root mean square threshold
$\aleph$	optimum TDNN model

Chapter 1

Introduction

Cough is common among children and can provide vital information on their airways. In this chapter, I describe the significance of cough analysis in diagnosing respiratory diseases in children, identify open research problems, and describe the research objectives and the contributions of this thesis.

1.1. Background to the research

Cough is one of the top reasons for visits to physicians in the world. It contributes to around 30 million and 1.4 million of total visits in United States and Australia, respectively [1, 2]. It is estimated that there are around 50 million cases of whooping cough alone per year, resulting in 300,000 child mortalities annually around the world [3]. Persistent cough is associated with deprivation of quality of life [4, 5], as well as anxiety and depression [6, 7]. Consequently, cough imposes a substantial economic burden due to medical consultations and medication [8-10].

Cough is a defensive mechanism that aims to protect the respiratory system [11, 12]. It can be stimulated by accidental events such as foreign body inhalation or by mucus produced internally. Cough is also one of the main symptoms of respiratory disease [13], ranging from mild colds to pneumonia. It is assessed by physicians when handling patients with respiratory complaints [14, 15], who seek information about its characteristics, quantity, and duration.

Cough carries substantial information on the health of the airways and is significantly helpful to determine the aetiology of disease, as shown in several past research studies [16-19]. Korpás, Sadlonová, and Vrabec [16] reported that cough sounds from patients with airways inflammations were dominated by lower frequency components (peak frequency ≈ 250 Hz) compared to normal patients (peak frequency ≈ 600 Hz). Smith et al. [17] and Hiew et al. [18] showed that wheezing was prominent in asthma coughs. In their study, Knocikova et al. [19] attempted to use cough to separate respiratory diseases. They claimed the 85 to 90 percent of correct rates on classifying voluntary coughs from healthy, asthmatic and chronic obstructive pulmonary diseases (COPD)

patients using wavelet coefficients [19]. The results show the potential of cough to be developed as screening tools.

Quantitative analysis of cough is not a mature field, particularly as clinicians have access to a limited number of basic tools with which to analyse cough. One of the main methods available to quantify cough is to count its frequency, that is, the number of coughs occurring in a given time unit [20, 21]. Knowing this information is useful to determine the severity of the cough and to assess the efficacy of treatments or new antitussive trials [22, 23]. The studies in [24, 25] reported that an increase in cough frequency correlated with inflammation of the airways. Long term monitoring of cough quantity enables the observation of the temporal pattern and duration of cough. The temporal information can be used to define the origin of cough. For example, high quantity of cough at night may indicate nocturnal asthma [26, 27], while increased cough after a meal may represent gastroesophageal reflux (GERD) [28, 29]. The duration of the coughs is also useful for diagnosis. Acute cough (cough lasting < 3 weeks) is associated with upper respiratory tract infections (URTI) [30], while chronic cough (cough lasting > 8 weeks) suggests lower respiratory tract infections (LRTI) [31].

Several efforts have been made to develop automated cough counting devices [32-36]. However, none of these devices has been tested in the paediatric population; the majority of the devices were tested using voluntary coughs from adults. Recording coughs in the paediatric population poses greater challenges. Young children, unlike adults, are not capable of following instructions in order to reduce activities such as crying that may interfere with the cough signals. Moreover, they have difficulty coughing on demand. Further, contact sensors, which were used by the majority of these devices, may not suit application to younger children for comfort reasons. Contact sensors also require an intensive sterilization process, especially when used with infectious diseases. Intensity/waveform-shape based methods for detecting cough used by previous studies are unlikely to be optimal when implemented in children with respiratory diseases such as pneumonia. In these patients, the intensity/waveform-shape of the coughs can vary radically. Moreover, the physiology and origin of cough in children are different from adults [37] and, as such, methods developed for adults cannot be directly applied to children.

The cough quantity can be observed manually [20]. The physician may listen to several episodes of cough or interview the patients and/or caretakers to obtain descriptions of the cough, as well as its influences on the quality of their life. Tools such as questionnaires [38-41], diaries [42], and scales [43] are used to document the long term information. However, the results of these measurements are subjective and depend on the skills, experiences, and vigilance of the observers [44-46]. Further, manual cough assessment is time consuming and a labour intensive process,

especially when working with long term data. Thus, there is an enormous need to facilitate automation on long time monitoring of chronic cough as well as acoustic studies of cough sounds, especially targeting the paediatric population.

One of the highly useful cough characteristics in paediatric cough is wetness/dryness [47]. Wet cough indicates the presence of mucus. In children, it is more likely correlated with lower respiratory tract syndromes and neutrophilic inflammation as a response to bacterial infection [48]. It is considered a significant marker to serious diseases such as pneumonia, asthma, bronchiectasis, chronic bronchitis, and bronchiolitis [49]. In contrast, dry cough raises minimal concerns since it represents less serious underlying conditions. It is associated with habit, GERD, upper airway infections, post infections, or hypersensitivity of cough receptors [31, 50-52]. Technology for automated classification of cough into wet/dry classes can be highly useful, if available.

Cough analysis can also help with differentially diagnosing respiratory diseases such as pneumonia [48]. Cough is a major symptom of pneumonia, a leading disease that contributes to around 1.3 million child deaths per year over the globe [53, 54]. The majority of pneumonia cases occur in low income countries where even the most basic diagnostic tools for diagnosing pneumonia such as x-ray are extremely rare. To manage childhood diseases including pneumonia, the World Health Organization WHO) has developed a guideline on Integrated Management of Childhood Illnesses (IMCI) [55]. Unfortunately, the IMCI guideline has poor specificity [56-58]. Other diseases such as asthma are often misdiagnosed as pneumonia, leading to unnecessary antibiotic treatments. This occurs because the symptoms used to screen in pneumonia (cough and rapid breathing) coexist in asthma [59, 60]. A recent study in Uganda [59] showed that 95 percent of 253 children with asthma syndrome received antibiotic treatments meant for pneumonia. These results are in line with a study in [61] where 46 percent of 200 children diagnosed with pneumonia (according to WHO criteria) actually had asthma. The excessive usage of antibiotics triggers bacterial resistance, and contributes to treatment failure rate as high as 22 percent in some regions [62]. Problems with differentiating pneumonia from asthma or other lower respiratory tract infections that do not need antibiotics have become a serious issue.

Researchers have attempted to improve the specificity IMCI by augmenting extra symptoms, such as fever, nasal flaring, chest in-drawing, poor sleep, or cough lasting more than two days [63-66]. There were improvements in specificities ranging from 20 to 90 percent, but these were at the cost of significant decline in sensitivity. The augmentation of symptoms also increases the complexity of the pneumonia diagnosis algorithm requiring trained clinicians for the implementation. In low-income countries, the lack of resources makes these approaches prohibitive to complete.

The updated WHO guideline recommends the bronchodilator test to separate asthma from pneumonia [67, 68]. Nevertheless, bronchodilators and their delivery systems (nebulizers and inhalers) are expensive, time consuming, and rare in remote areas. An alternative method for the bronchodilator test is urgently required.

In this thesis, I address these issues while targeting the paediatric population. The methods developed in this thesis are intended to address some of the fundamental gaps in the field of cough sound analysis.

Details of the objective of this thesis are presented in the following section.

1.2. Research problems and objectives

The research objectives of this thesis are as follows:

1. **To develop an automated method to classify cough into wet/dry classes.**

   Currently, to determine the wetness/dryness of the coughs, paediatricians manually listen to several voluntary/non-voluntary coughs. The results of this manual classification are subjective, dependent on the skill and experience of the paediatricians. In this thesis, an automated technology is developed to classify cough into wet/dry classes.

2. **To develop a cough sound based method for separating paediatric pneumonia from asthma.**

   Separating pneumonia from asthma is a major problem in remote areas. The WHO guidelines call for a bronchodilator test, which is expensive and time consuming to administer in resource poor regions. A cough sound analysis based method is proposed in this thesis as a substitute for bronchodilators. This approach is unique and advantageous, as cough signal can be acquired using low cost, non-contact recorders.

3. **To develop a method to automatically segment cough sounds from recordings obtained in paediatric wards.**

   Cough events can be obtained manually by listening to the recordings. However, the process is tedious and costly, especially when working with a large number of long duration recordings. In this thesis, an automated technique is developed to segment cough sounds from a continuous sound recording.

1.3. Organization of the thesis

The remainder of the thesis is organized as follows:

Chapter 2: This chapter describes the respiratory system and respiratory diseases and provides a key section on paediatric pneumonia. Cough sound analysis is then presented to provide the foundation of this thesis. The chapter concludes with a brief discussion on the potential of cough as a diagnostic tool.

Chapter 3: This chapter describes the data acquisition systems, protocol, inclusion, and data pre-processing used in this thesis.

Chapter 4: In this chapter, the development of automatic wet/dry cough classification in paediatric populations is described. The method is intended to facilitate objective wet/dry cough identification.

Chapter 5: This chapter describes the development of a novel method for pneumonia/asthma classification. The method is proposed as the substitute for the bronchodilator test in resource-limited settings.

Chapter 6: This chapter explains the development of automatic cough segmentation for paediatric populations. A combination of sound features is proposed to avoid dependency on intensity based features. The method is intended to address manual cough counting in the paediatric population, as well as front end of cough analysis system.

Chapter 7: In this chapter, the studies presented in thesis are discussed, followed by suggestions for future research and a conclusion to the thesis.

Bibliography

Appendices

1.4. Contribution of the thesis

The contributions of this thesis to the field of research consist of the development of novel methods for objective cough analysis. The contributions of this thesis are as follows:

- *Development of cough sound analysis methods in paediatric populations*. All coughs used in this study were spontaneously obtained with non-contact sensors. This study is novel, as most existing studies on cough sound analysis include adult subjects with voluntary coughs. Methods developed for adults are not feasible with paediatric populations.
- *Development of a novel method for automated wet/dry coughs classification*. This study is the first in this field, opening a new branch of research on cough analysis in paediatric populations. The method achieved high classification accuracy when compared to the work of two paediatricians, each of whom had more than 15 years of clinical experience. The method facilitates the objective classification of a wet/dry cough study over a longer period.

- *Development of a novel method for discriminating pneumonia from asthma using cough sound analysis*. The method can be used to support the existing WHO guideline in pneumonia management as a substitute for the bronchodilator test in resource-limited settings.
- *Development of an automated cough segmentation algorithm targeting paediatric populations*. One of the novelties of the research is that the proposed method utilizes a combination of features instead of depending solely on the magnitude of the cough signal. This approach is beneficial in the paediatric population in terms of respiratory disease where the magnitude of cough sounds can vary widely. The method can be implemented as cough counting devices as well as the front end of cough sound analysis.

An overview of the proposed methods is illustrated in Figure 1.1.

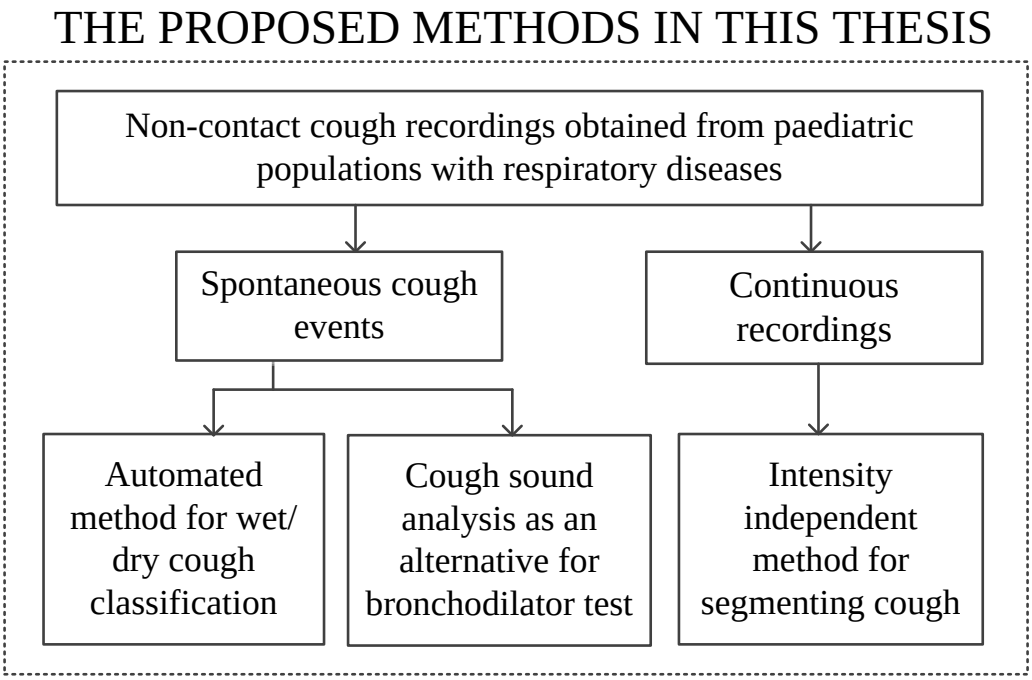

Figure 1.1: Overview of the methods carried in this thesis

Chapter 2

Respiratory Systems and Cough Sound Analysis Based Diagnostics

Cough is one of the indications of respiratory disorders. In this chapter, I describe the overarching view of the respiratory system followed by a description of respiratory system infections. The description of cough and its diagnostic values are then presented to provide the foundation for this thesis.

2.1. Respiratory system

Respiration/breathing is one of the basic life functions of living organisms. In humans, the process is responsible for maintaining the exchange of oxygen and carbon dioxide, as well as to regulate blood acidity [69]. The breathing process is comprised of two processes: inspiration/inhalation and expiration/exhalation.

During inspiration, the skeletal muscles of the diaphragm, abdomen and rib cage contract, causing the chest to expand as air rich in oxygen flows into the nose, nasal cavity, pharynx, and larynx [70]. Air can also pass through the mouth instead of the nose. All these organs are usually called upper respiratory tract organs [71]. The nasal cavity serves several important functions. It has hairs and mucosa to filter airborne particles and to trap bacteria [72]. Cilia slowly move the trapped particles/bacteria in the mucosa to the pharynx, where it is swallowed and digested in stomach. This process protects the lungs from inhaled bacteria and foreign materials. The nasal cavity also adjusts the air to one's body temperature and humidifies it [70]. In the larynx, there are vocal folds and a glottis for speech/sound production. The glottis also functions as a lid for larynx to protect the lungs from foods when swallowing [72].

From the larynx, the air passes through lower respiratory tract organs comprises of trachea, bronchi, and lungs [71]. A mucous membrane with hair-like cilia exists on the inner surface of trachea [72]. The cilia constantly clean the tract and carry foreign substances upward for swallowing or expectoration. As the conducting airways, trachea branch into the right and left main bronchi at the level of the sterna angle [73]. The main bronchi branches into the lobar bronchi and

then the segmental bronchi. The segmental bronchi continue to branch into the smallest bronchi called bronchioles. The bronchioles branch to terminal bronchioles, then to alveolar ducts, and finally terminate in a small collection of air sacs known as alveoli [73]. Alveoli have a thin, single layer of epithelium cells covered by a cobweb of pulmonary capillaries. This structure facilitates oxygen and carbon dioxide exchange through the diffusion process [72, 73].

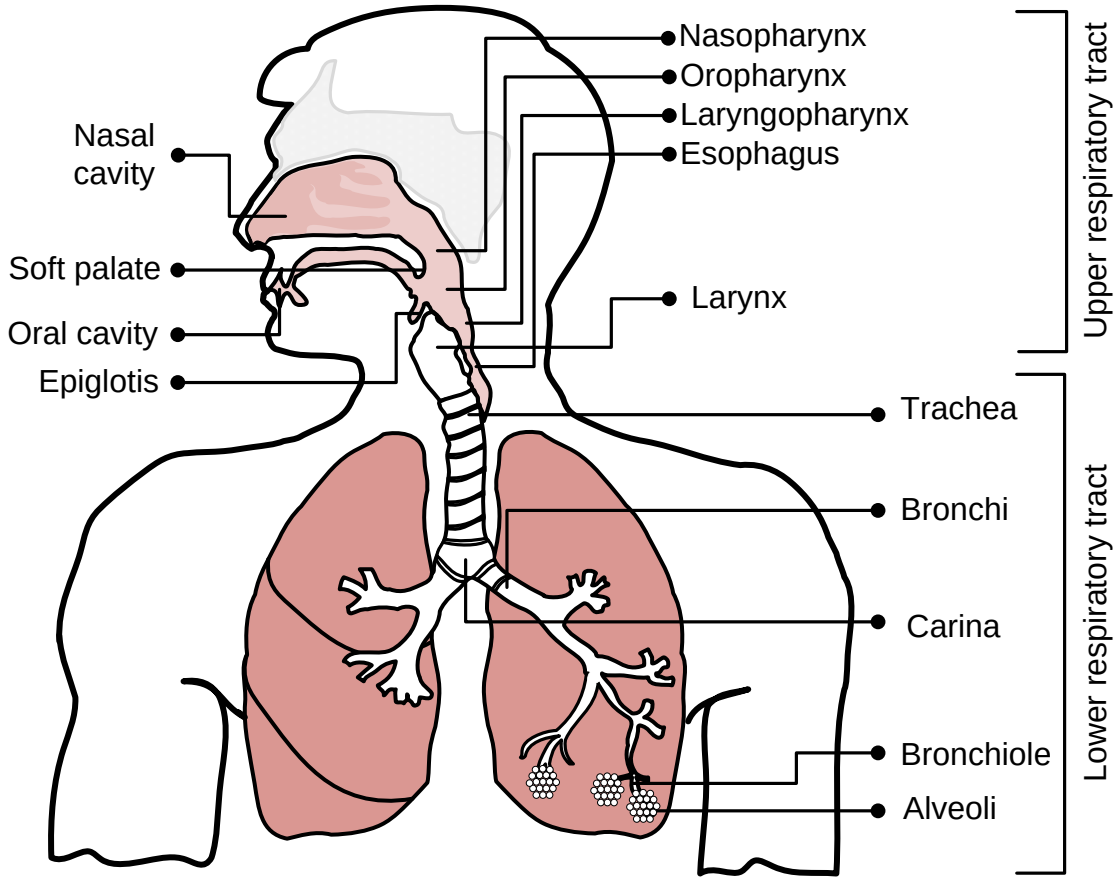

Figure 2.1: Human respiratory system

Expiration occurs when the skeletal muscles of diaphragm, abdomen, and rib cage are relaxed [70], causing the chest to depress and the lung volume to decrease; as such, air containing carbon dioxide flows out in the opposite direction to the inspiration. Normally, expiration is a passive-effortless activity [72]. However, expiratory can be active (forced expiration), where the muscles depress the rib cage, abdomen, and diaphragm to force air flows out from the lungs. This condition may occur due to respiratory system infections such as asthma where the airways are narrowed by spasms of the bronchioles; it can also occur in the case of pneumonia where excessive mucus fills the alveoli [74].

The human respiratory system is illustrated in Fig 2.1. The right lung has three lobes and is bigger than left lung which only has two lobes [75]. The lungs are surrounded closely by two membranes known as the visceral pleura, which attaches to the lungs, and the parietal pleura, which attaches to the chest wall [76]. A thin layer of fluid presents between the two layers of pleura to lubricate the membranes and facilitate the frictionless motion of the lungs against the chest wall during breathing. The fluid also provides molecular cohesive forces that prevent the separation of the lungs from the chest wall under normal circumstances. This allows the lungs to follow the direction of the chest wall and diaphragm with breathing.

2.2. Respiratory diseases

Respiratory system infection is a medical term referring to the impairment of the respiratory organs. Respiratory diseases can be caused by bacteria, viruses, fungi, pollutants, or genetic problems [77]. Based on the division of respiratory systems, there are two major respiratory system infections called upper respiratory tract infection (URTI) and lower respiratory tract infection (LRTI) [78]. URTI consists of common diseases in the population, including influenza, the common cold, pharyngitis, epiglottitis, sinusitis, and laryngitis; fortunately, the majority of diseases are self-limited and not life threatening.

Normally, the lower respiratory tract below the larynx is sterile. However, infections can reach these parts by inhalation, aspiration, direct inoculation, or they can be blood borne [78]. The manifestations of LRTI include pneumonia, asthma, bronchitis, tuberculosis, or lung cancers.

The main symptoms of respiratory system infections are widely varied, and include cough, rapid breathing, dyspnoea (shortness of breath), wheezing, haemoptysis, and chest pain [79, 80]. Associated symptoms that may accompany the main symptoms are anorexia, weight loss, pyrexia, and sweating. Hypoxemia is indicated by lethargy, malaise and confusion. Headaches, particularly on awakening in the morning, may represent the symptom of hypercapnia (a high level of carbon dioxide level in the blood) [81]. Oedema may indicate pulmonary heart disease, while snoring and daytime somnolence may indicate obstructive sleep apnoea syndrome. Voice hoarseness is usually manifested as inflammation in the vocal cords or laryngitis [81, 82].

2.3. Paediatric pneumonia

Pneumonia is a serious lower tract respiratory system infection and highly prevalent in the paediatric population [83]. This is the main disease contributing to the high mortality rate in the paediatric population, especially in children younger than five years old who live in low income

countries [83]. In this century, even a basic diagnostic tool such as x-ray is rarely available in the top level health facilities in these countries to establish the correct diagnosis. Thus, problems in pneumonia diagnosis have become a serious issue and it is crucial that this be addressed.

This section will provide the background about pneumonia and pneumonia diagnosis problems as well as a discussion on cough sound analysis. The last sub-section summarises the preliminary works achievement, the gaps in the cough analysis field, and the proposed studies.

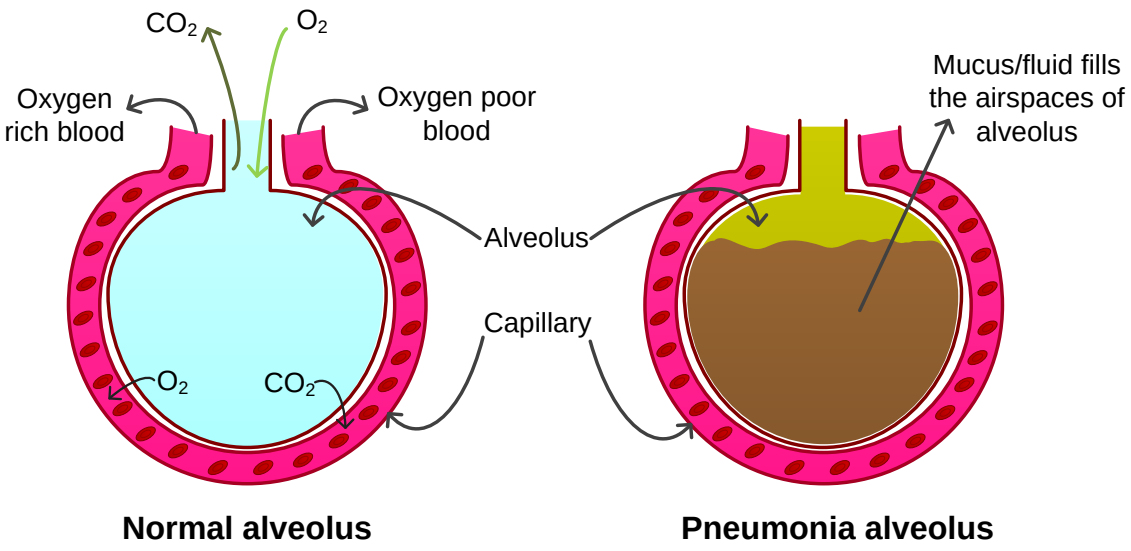

Figure 2.2: Illustration of alveolus in the normal condition and pneumonia. In pneumonia, oxygen is unable to reach the blood stream due to the mucus in the alveolus.

2.3.1. Definition of pneumonia and its consequences

Pneumonia is defined as an acute infection/inflammatory disorder of the lung parenchyma (alveoli and area between alveoli/interstitial) [48]. The inflammation excites effusion where mucus (white blood cells, red blood cells, and fibrin) pour into the alveoli (see Figure 2.2). In the case of serious inflammation, the mucus may completely fill the alveoli; this condition is called alveolar consolidation [84]. The existence of mucus in the alveoli reduces the alveoli's functionality to absorb oxygen and release the carbon dioxide, which causes hypoxemia (an abnormally low level of oxygen) and acidosis (increased acidity in blood) [84]. Inaccurate diagnosis and late treatment of pneumonia can lead to fatal consequences.

Pneumonia is the major cause of morbidity and mortality in paediatric populations. A study by Walker et al. [53] showed that in 2010 there were around 120 million episodes of pneumonia in children younger than five years. It was estimated around 1.3 million children died due to

pneumonia in 2011 [53]. Around 97 percent of pneumonia cases occurred in low income countries and 74 percent of the cases occurred in South Asia and sub-Saharan regions [85].

Paediatric pneumonia cases in industrialized countries are lower compared to low-income countries. The estimation of pneumonia incidence in European and American regions were 0.06 and 0.10 episodes per child-year, respectively. In these regions, pneumonia causes for around 12 percent of children deaths [85].

Pneumonia contributes to significant economic spending; annual costs incurred as a result of pneumonia were estimated at around €10.1 billion in Europe, $10 billion in the US, and $63 million in New Zealand [48, 86, 87]. The Australian Lung Foundation estimated that pneumonia created a cost burden of around $300-350 million a year in Australia [88]. The economic loss due to pneumonia in low income countries is unclear. However, the cost estimation for pneumonia treatment in these countries is enormous, as seen in Figure 2.3.

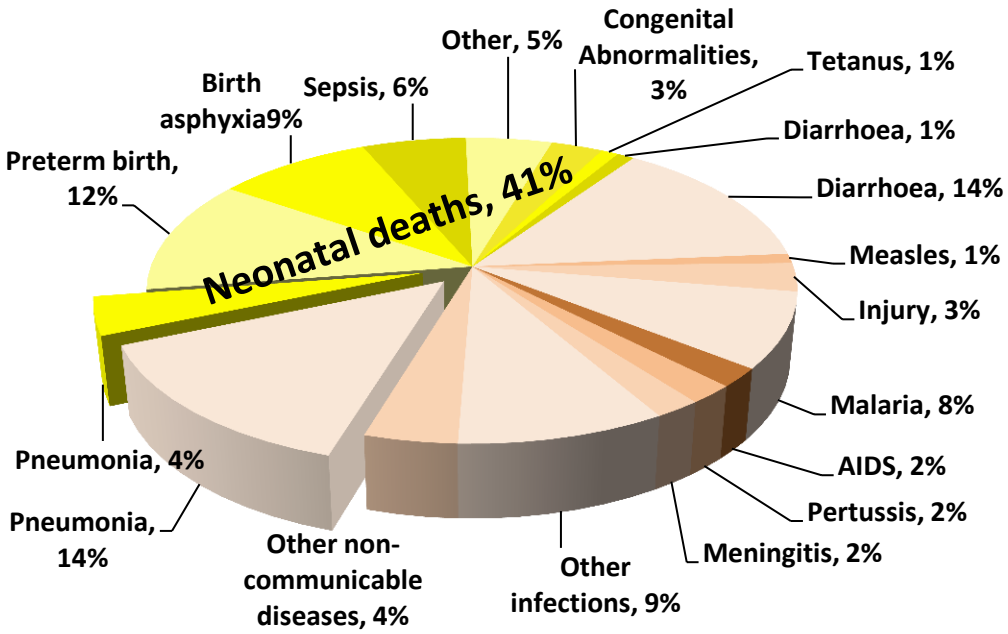

Figure 2.3: Diseases contributing to the mortality of children less than 5 years of age. Pneumonia is the leading disease, accounting for 18 percent of total child deaths.

2.3.2. The aetiology and epidemiology of pneumonia

Pneumonia can be classified according to one of three pathogen types: i) bacterial pneumonia; ii) viral pneumonia; and iii) atypical pneumonia. Bacteria are the major cause of hospital-acquired pneumonia in children, accounting for around 50 percent of cases [89]. Streptococcus pneumoniae is the most common pathogen causing bacterial pneumonia in infants and children [89, 90] and

accounts for around 16 to 37 percent of the cases. Other bacteria that may cause pneumonia include Staphylococcus aureus, Klebsiella pneumonia, Legionella species, and Moraxella catarhalis.

Respiratory syncytial virus (RSV) is the main pathogen that causes viral pneumonias. It contributes to around 14 to 40 percent of viral pneumonias [89]. Other types of viruses such as Parainfluenza, Influenza, and Adenovirus account for 15 to 20 percent, 8 to 27 percent, and 5 to 9 percent of viral pneumonias [89] respectively.

Atypical pneumonia refers to pneumonia caused by Mycoplasma pneumonia and Chlamydia pneumoniae and other rarer bacteria such as Chlamydia psittacci, Legionella pneumophilia and Coxiella burnetti [89]. Atypical pneumonia is commonly found in children older than five years [91]. Mycoplasma pneumonia accounted for 7 to 30 percent and 14 to 51 percent of atypical pneumonia in children aged 5 to 9 years and 10 to 16 years respectively. In the same age groups, the number of atypical pneumonia cases caused by Chlamydia pneumoniae was lower, ranging from 9 to 13 percent and 14 to 35 percent, respectively [92, 93].

2.3.3. Risk factors

The prevalence of paediatric pneumonia increases in lower socioeconomic groups. Pneumonia correlates with malnutrition, low birth weight, and a lack of breastfeeding [94, 95]. The risk of pneumonia infections escalates in children with underlying diseases such as immunodeficiency disorders, measles, congenital heart diseases, neuromuscular disorder, and gastrointestinal disorder [96]. Environmental factors such as crowded homes, cigarette smoke, and indoor air pollution from cooking activities also increase the risk of pneumonia [97].

2.3.4. Symptom and diagnosis

The diagnosis of pneumonia can be established by determining the symptoms and the aetiology. Cough and/or high breathing rate are the common symptoms of pneumonia [55]. Other symptoms that may appear in pneumonia are high heart rate (tachycardia), malaise, chills, and fever as a response to infection [98]. Severe pneumonia is indicated by the existence of lower chest in-drawing and blueness in skin due to hypoxemia [48].

Crackles, defined as a discrete explosive non-musical discontinuous abnormal lung sounds, can be found through the auscultation of pneumonia subjects [99, 100]. The sounds are produced by the sudden opening or closing of collapsed airways and the movement of excessive mucus [101]. Wheezes, a continuous coarse whistling sound produced by the narrowing of the respiratory tracts, are likely to occur in viral pneumonia [102].

Chest x-ray is used to confirm consolidation due to pneumonia. The consolidation is indicated by an area of increased density (grey/white area) involving a small lung segment, a lobe, or a whole lung [84, 98]. Air bronchogram, defined as the phenomenon of air-filled bronchi being made visible by the opacification of surrounding alveoli, may appear when the consolidation intensifies [103]. The aetiology of pneumonia can be determined by bio-chemical testing of the samples from the subjects such as lung fluid/tissues, blood, sputum, or urine [48].

2.3.5. Challenges in pneumonia management

Pneumonia is a dangerous threat for children, especially those who live in low income regions with limited access to health resources [104]. Fortunately, pneumonia is preventable and curable and there are several international organizations involved to fight this disease. The World Health Organization (WHO) and the United Nation International Children's Emergency Fund (UNICEF) have developed programs for pneumonia management including raising awareness of pneumonia, preventing pneumonia, and diagnosing and treating pneumonia. In 2013, WHO and UNICEF initiated a new program called the integrated Global Action Plan for Pneumonia and Diarrhoea (GAPPD). One of the targets of the program is to reduce the victims of pneumonia to zero by 2025. Also participating in this effort are the Bill and Melinda Gates Foundation and PATH, two organizations that encourage innovations for vaccinations, medication, and diagnostic tools to make these resources accessible for people in remote areas.

The invention of new pneumonia diagnostics tools is one of the great challenges in the management of paediatric pneumonia [105]. Currently, the diagnosis of pneumonia in resource-limited settings relies on the guidelines developed by WHO and UNICEF, called the Integrated Management of Childhood Illness (IMCI) [55]. The implementation of IMCI revealed that asthma cases are often misdiagnosed as pneumonia [59, 60]. The updated WHO guideline recommends the bronchodilator test to separate asthma from pneumonia; however, the test is costly, time consuming and rarely available in remote areas. There is an urgent need to develop a novel method to substitute for the bronchodilator test in remote areas.

2.4. A frame work for cough sound analysis

Cough is a defence mechanism used by the body to clear the respiratory tract of foreign materials inhaled accidentally or produced internally by infections [13]. It can be produced voluntary but is usually involuntary by reflex. Cough can be stimulated by: chemical compounds (e.g. cigarette smoke); foreign body inhalation (e.g. food); organ defects (e.g. vascular

ring/congenital defect); temperature (e.g. cold dry air); and inflammation (e.g. excessive mucus in the airways due to respiratory system infections). In some cases, irritation of the airways without mucus can stimulate cough as well. Hyper-reactivity of irritant receptors on the respiratory tract mucosa due to inflammation or other pathologic process may enhance the cough reflex to react with mild irritation [98]. Cough can also be stimulated by irritation in the pleura and tympanic membranes [106].

The European Respiratory Journal (ERS) guidelines recommends two definitions of cough [107]: first, as a "three-phase expulsive motor act characterized by an inspiratory effort (inspiratory phase), followed by a forced expiratory effort against a closed glottis (compressive phase) and then by an opening of glottis and rapid expiratory airflow (expulsive phase) [108]; and second, as "a forced expulsive manoeuvre or manoeuvres against a closed glottis and associated with a characteristic sound or sounds".

Based on physiological considerations, Hotaling and Moynihan [13] defined cough as having four different phases: inspiratory, contractive, compressive, and expulsive. The inspiratory phase is initiated by breathing in and is terminated by the closure of the glottis. In the contractive phase, groups of respiratory muscles contract, leading to a marked elevation of alveolar, pleural, and subglottic airway pressures. In the expulsive phase, the glottis opens quickly, followed by a rapid exhalation of air under a large pressure gradient. The rapid movement of air expelled from the lung generates the cough sounds with contributions coming from different areas of the respiratory system, such as the narrowing airways, the vocal fold, and the mucus vibration [98, 109]. The cough sound, therefore, carries information about the nature of the infections. Moreover, in paediatric populations, the characteristics of cough suggest the aetiology of the diseases [110, 111].

Previous studies have examined the acoustics of cough sounds [16, 112-114]. However, none of those studies explored the cough in paediatric populations. They focused heavily on the classification of cough and non-cough sounds [112-113], the acoustic sound analysis of cough from healthy subjects and ill subjects [16, 114], and on counting the number of coughs [32-36]. Only two studies [19, 115] attempted to explore the characteristics of cough sounds or to differentiate cough in different respiratory system infections. Thus, the potential of cough to be developed as a diagnostic tool has not been investigated thoroughly.

2.5. Summary

The physiology of respiratory systems change as a result of respiratory system infections. Infection stimulates inflammation and the hyper-secretion of mucus in the airways or respiratory

organs and contributes to the wetness of the coughs. The wet/dry cough information is useful for differential diagnosis. In the paediatric population, it may indicate lower respiratory tract infection [110]. Currently, however, as there is no available method to identify wet/dry cough automatically, physicians must identify wet/dry cough manually. The results from the manually assessed wet/dry cough is subjective depend on the skills and experience of the physician. A novel method to automate the wet/dry cough identification process is required.

In lower respiratory tract infections such as pneumonia, excessive mucus/fluid fills the alveoli and causes lung consolidation. The fluid filled alveoli reduce the normal capacity of the lung in breathing. Further, the sudden opening or closing of the alveoli collapsed by mucus creates crackle sounds. The lung consolidation, the reduced capacity of the lungs, the vibration of the mucus, and the crackles themselves all work together to alter the acoustics of pneumonia cough sounds. In contrast, the acoustics of cough in asthma is mostly affected by narrowed airways due to the spasm of bronchioles. This shows that the acoustics of cough carries information related to respiratory tract infections. The acoustic features of cough sounds can be extracted using signal processing and used to differentiate asthma from pneumonia.

Despite the significance of cough in the diagnosis, the quantitative analysis of cough is still immature. Existing studies focus heavily on the development of cough counting devices. To obtain cough samples for analysis, health practitioners and researchers must segment the samples manually (determine the position, start and end of cough events), which is both tedious and time consuming. Therefore, an automated cough segmentation method is urgently needed.

In this thesis, I address the problems of wet/dry cough classification, paediatric pneumonia/asthma separation, and cough segmentation from the recordings. I describe the development of the methods in Chapters 4, 5, and 6, respectively. In the following chapter, I describe the data acquisition techniques used in this thesis.

Chapter 3

Research Data Acquisition

This chapter describes the clinical data acquisition systems and procedures for cough recordings in paediatric population. To the best of my knowledge, this study is the first effort to analyse the acoustics of non-voluntary cough sounds in paediatric populations with respiratory diseases. As such, a typical database is not yet available in the public domain. In this thesis, data acquisition is a substantial part of the work. The aim of data acquisition is to record high quality cough sounds such that all the sound features can be preserved. Data from the data acquisition were used to develop the methods of this thesis.

3.1. Data acquisition systems

The cough recording system consisted of a low-noise microphone with a cardioids beam pattern (Model NT3, Rode®, Sydney, Australia), followed by a pre-amplifier and an A/D converter (Model Mobile Pre-USB, M-Audio®, CA, USA). The output of the Mobile Pre-USB was connected to the USB port of a laptop computer. The nominal distance from the microphones to the subjects' mouths was 50 cm with a possible variation in actual distance from 40 to 100 cm due to subject movement. The sampling rate was set at fs = 44.1 k samples/s with a 16-bit resolution to obtain the best sound quality. The data was acquired in the natural hospital environment, without modifying it in any way, other than placing the sound recording system by the bed (see Figure 3.1). The duration of recording for each subject varied from 1 to 6 hours, depending on the situation or condition of the patients.

3.2. Protocols

The data for this work were recorded at Sardjito Hospital, Yogyakarta, Indonesia, from paediatric patients admitted with respiratory complaints. Table 3.1 shows the inclusion and exclusion criterion for the subject's recruitment. The inclusion criteria used in the recruitment was patients with at least two of the following symptoms: cough, sputum, breathlessness, and temperature higher than 37.5°C. I excluded patients with advanced disease where recovery is not

expected, such as those with lung cancer, or diseases with droplet precautions such as tuberculosis, and patients undergoing ventilation treatment.

The recordings were started after paediatricians had examined the subjects, initial treatment had begun, and informed consent had been completed. The diagnoses of disease were established based on clinical examination and were supported by laboratory test results (e.g. chest x-ray, blood test, C-reactive protein (CRP), etc.). The research protocol for this study received ethics clearances from Sardjito Hospital and The University of Queensland, Australia.

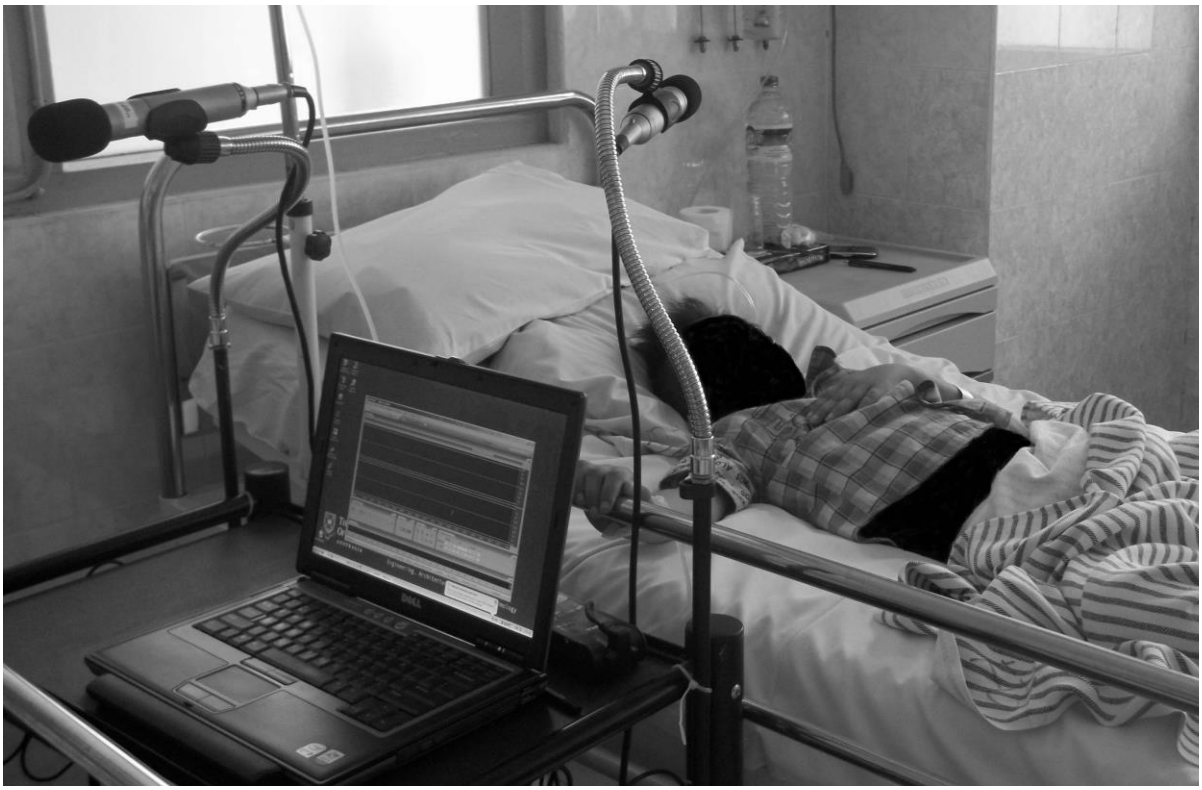

Figure 3.1: Data acquisition systems set up

Table 3.1: Inclusion and exclusion criteria used in the study.

Inclusion criteria	Exclusion criteria
- Symptoms of respiratory tract infection , at least 2 of the following: cough, sputum, breathlessness, and temperature >37.5° - Informed consent complete	- Advanced disease where recovery is not expected. - Droplet precautions - Having non-invasive ventilation treatment - No informed consent

3.3. Database

There are two types of data obtained from the data acquisition: first, cough sound recordings saved as regular wave files (*.wav); and, second, clinical diagnostic information pertinent to the study. In all circumstances, the data is de-identified by removing the identification details (e.g. name, address) before transferring it from the hospital site.

In each recording, cough events were manually identified and marked. To define the beginning and end of cough segments, the scorer listened carefully to the sounds and simultaneously looked at the time domain waveform displayed on the computer screen. The criterion of the cough events selections were: first, the cough signals did not overlap with other sounds; and, second, the cough signals were not clipped. To listen to and display the sound signal I used Adobe Audition software.

3.4. Data modelling and pre-processing

The sound recordings obtained from the data acquisition process contain several components. In this thesis, the sound recordings are modelled in mathematic equations to describe their components. The mathematical equations are also used to illustrate the changes in the sound recording components after the application of signal processing. The notations of the mathematic equations are used consistently in the development of the methods in thesis.

Let $s[n]$ be the discrete time sound recording. The expression of the signal $s[n]$ is given as follows:

$$s[n] = s_y[n] + b_1[n] + b_2[n] \tag{3-1}$$

where $s_y[n]$ is the audio sound signals, $b_1[n]$ represents the low frequency noise in the measurement (e.g. noise coming from the vibration of microphone's stands/leads) while $b_2[n]$ is the Gaussian noise (mostly high frequency). The audio signal s_y[n] is comprised of cough sound signals $s_{cg}[n]$ and non-cough sound signals $s_{nc}[n]$, as given in (3-2).

$$s_y[n] = s_{cg}[n] + s_{nc}[n] \tag{3-2}$$

To reduce the noise, $s[n]$ was processed through two different filters: a High Pass Filter (HPF), and a Power Spectral Subtractions (PSS) filter. The HPF was implemented in the first stage to reduce the low frequency noise $b_1[n]$. It was designed with a cut-off frequency of f_c = 10 Hz. In the second stage, the PSS filter was employed to reduce the Gaussian noise $b_2[n]$ [116]. After the HPF and PSS filters, the estimate $\hat{s}[n]$ of the recording $s[n]$ is given by:

$$\hat{s}[n] = \hat{s}_y[n] \tag{3-3}$$

where $\hat{s}_y[n]$ represents the estimates of audio sound signals comprised of estimates of cough sounds $\hat{s}_{cg}[n]$ and non-cough sounds $\hat{s}_{nc}[n]$, respectively. Therefore, Equation (3-3) can be re-written as follows:

$$\hat{s}[n] = \hat{s}_{cg}[n] + \hat{s}_{nc}[n] \tag{3-4}$$

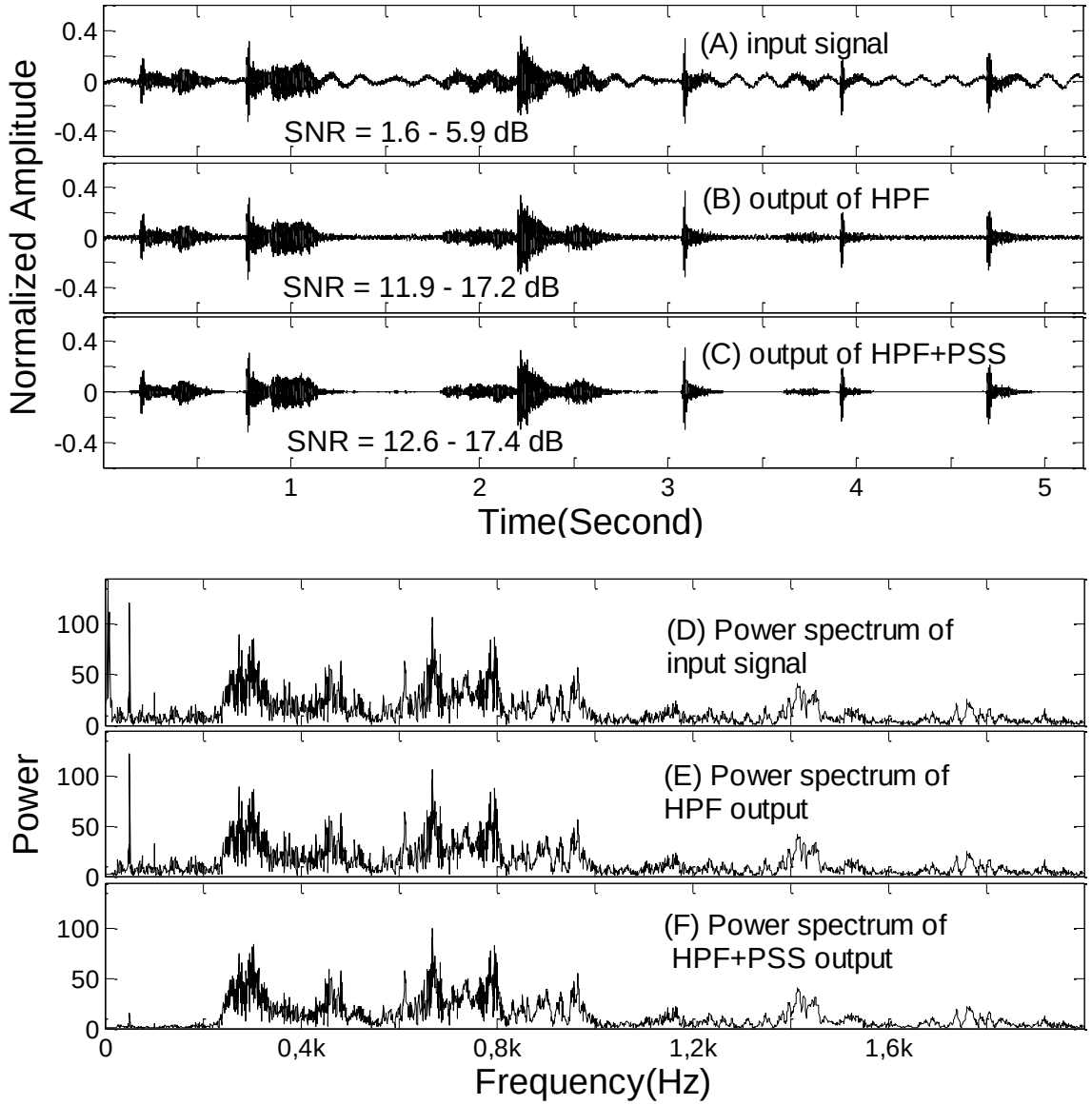

Figure 3.2: Illustration of the noise reduction process in time domain (figure A, B and C) and frequency domain (figure D, E and F). HPF = High Pass Filter and PSS = Power Spectral Subtraction filter.

Figure 3.2 illustrates the effect of HPF and PSS filters to remove the noise components from the raw signal. From the frequency spectra (Fig 3.2(D)) of the raw signal in Fig 3.2(A), it can be seen that the raw signal contains low frequency interference (< 10 Hz) as well as 50 Hz interference. The signal to noise ratio (SNR) is defined as SNR= 20log10(|Signal Magnitude|/|Noise Magnitude|) dB. From the figure, it can be seen that the SNR of the raw signal in this example is very poor (< 6 dB).

Figures 3.2(B) and 3.2(C) show the output signal $\hat{s}[n]$ after the application of HPF and PSS filters. Their frequency spectra are shown in Fig 3.2(E) and Fig 3.2(F), respectively. According to these figures, the designed filters reduce the low frequency interference (< 10 Hz) and 50 Hz interference significantly. The SNR of the filtered signal $\hat{s}[n]$ increased to > 17 dB.

In the next chapters, I describe the development of the method for wet/dry cough identification using the cough from the filtered sound signal $\hat{s}[n]$.

Chapter 4

Automatic Identification of Wet/Dry Cough in Paediatric Patients

Cough can be classified as wet or dry cough. This is a clinically useful classification in paediatric populations, where wet cough is more likely to be indicative of lower respiratory tract bacterial infections. Considerable experience is needed to classify coughs accurately into the two categories in actual clinical practice. In this chapter, I propose a method for classifying coughs into wet and dry classes. The method is objective, fully automated, and has the potential to be used in long term cough monitoring applications.

4.1. Introduction

Cough can be categorized as wet or dry, based on its acoustic quality. Cough is characterized as wet when the sounds carry features indicative of mucus; in the absence of perceivable wetness, they are called dry. Medically, there are different reasons for wet and dry coughs and their identification aids in the differential diagnosis of diseases such as bronchiectasis, asthma, chronic bronchitis and bronchiolitis [49]. Often, the wet/dry classification is used in epidemiological studies [117, 118] and clinical research [49, 119]. In children, wet cough is generally associated with lower respiratory tract bacterial infections [119]. Diseases such as asthma and post-infections can cause dry cough. In some cases, the presence of dry cough as perceived by a clinician indicates an early stage of the disease, which may later become wet cough as the disease progresses, leading to more secretions in airway.

Currently, cough quality is evaluated by clinicians either by asking the patient or the patient's caretakers to describe their cough during clinical assessment or by asking them to cough and listening to the cough. However, while doing so significant temporal information about the frequency of coughs and variation in wetness of the cough is lost, information that may be useful in making a differential diagnosis and assessing the efficacy of the treatment. In addition to this, the manual evaluation of the wetness of a cough is a subjective process and the outcome depends on the experience of clinicians [17, 120]. The process also suffers from the difficulties inherent in

discerning, via coughs, low-levels of mucus in airways; even trained clinicians underscore wet coughs as confirmed by bronchoscopic findings [49].

Researchers have rarely attempted to develop technology for the automated, objective classification of cough into wet/dry categories. To the best of my knowledge, only two prior works exist in this area [121, 122]. Murata [121] argued that cough sound frequencies can be used to discriminate between wet and dry coughs. Chatrzarrin et al. [122] proposed peaks of the energy envelop and spectral features of the cough sound for the same purpose. These studies opened up a new branch of research in respiratory sound analysis, despite being limited to a descriptive study of some characteristic features of coughs. No definitive classification algorithm or results were presented for wet/dry differentiation. The amount of data analysed was fairly limited as well, with only 30 cough samples from 10 subjects (5 healthy and 5 bronchitis patients) and a total of 16 coughs in their dataset, making the interpretation of the results difficult.

All previous research used cough sounds from adult subjects only and adopted techniques that used duration, magnitude, and frequency features to characterize cough into dry/wet categories. Cough in adults is different in many ways; while wet cough is the term used with children, 'productive cough' is the term used for adults, as they are able to expectorate airway secretions. Further, the same amount of secretions in a large airway (i.e. in adults) would biologically produce a different sound in a small airway (i.e. in children). Further, production of cough sound is a complex physiological process involving several anatomical structures in the upper and lower system. Its acoustic properties vary significantly [114] with individual differences, age, gender [123] and the state of the airways is a significant factor as well [16]. In disease, cough sound characteristics may change, making it necessary to develop robust methods to identify wetness/dryness. Intensity and duration dependent methods will not be sufficient to capture the rich information hidden in cough sounds.

Cough can be a symptom of serious diseases such as childhood pneumonia which kills more than a million children in the world [53]. The clinical community recognizes the important of cough in assessing the health of children. However, researchers have rarely attempted to develop objective, automated cough analysis systems for children. In particular, no prior work exists in the area of wet/dry classification. Cough assessment technology developed for adults cannot be extrapolated for children [37]. There is an urgent need to develop an automated objective cough assessment method for children.

In this chapter, these issues are addressed and an automated objective classification model is proposed to categorize cough sounds into wet and dry classes. The method uses first, second, and third order statistical features (e.g. formant frequencies, Mel-cepstrum, non-Gaussianity, and

bispectrum etc.) of the cough sounds. A linear regression model was used as classifier to differentiate wet cough from dry cough using the computed features.

This research on wet/dry cough identification involving paediatric populations is the first effort in this field. The method can be used to assists paediatricians to determine the cough characteristics objectively and has the potential to be developed for long term cough monitoring. It can be used to study the pattern of wet/dry coughs and to assess the effectiveness of a medication.

4.2. Material and method

Figure 4.1 shows the block diagram of the automated cough classification algorithm proposed in this study. It is divided into three stages: (A) the creation of a cough sound database and classification into wet/dry classes by an expert scorer; (B) the design of an automatic classifier; (C) testing the classifier on a prospective cough sound dataset. Details of the method are described in Sections 4.2.1 to 4.2.3.

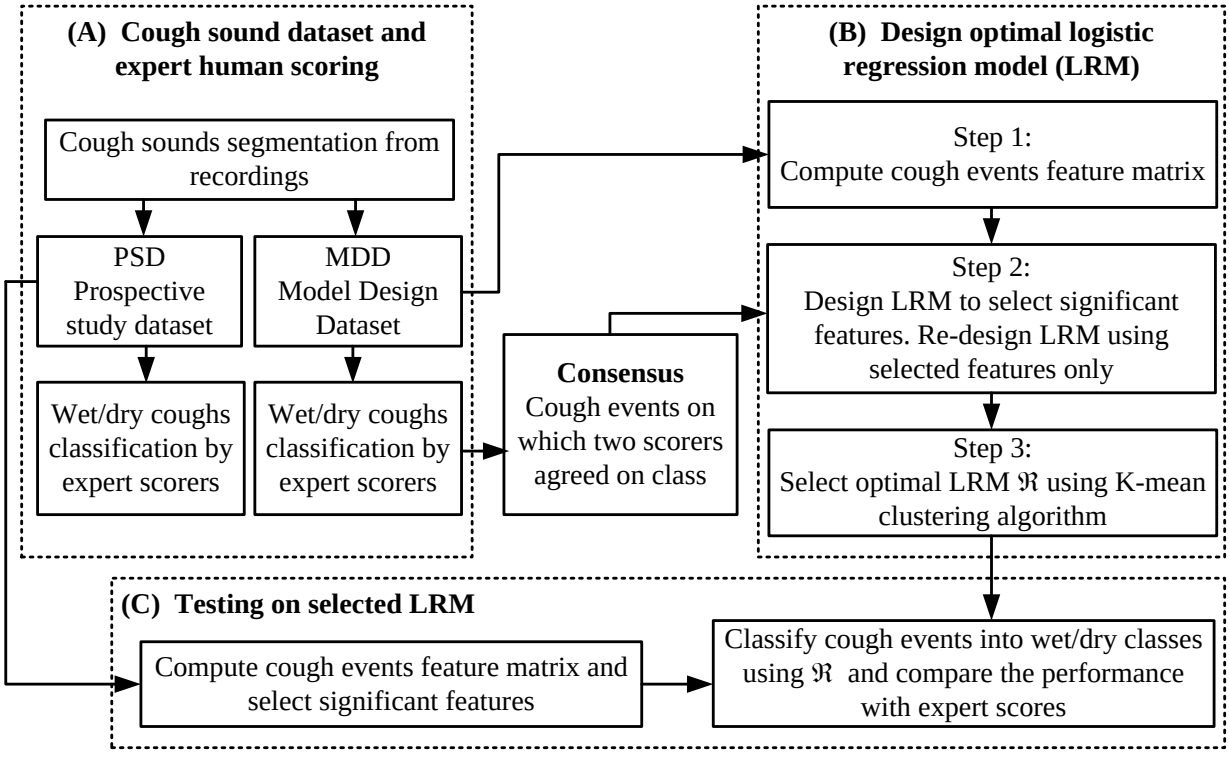

Figure 4.1: Block diagram of the proposed method for wet/dry cough sound classification.

4.2.1. Cough sound data and classification into wet or dry by expert human scorers

Let Ɒ be the number of patients whose sound recording is used in this study and Ç be the total number of cough events from Ɒ patients. The Ɒ patients with Ç cough events were divided into two datasets: first, MDD (model design dataset); and second, PSD (prospective study dataset). The patients were divided into MDD and PSD based on the order of presentation to the hospital respiratory clinic. Patients in datasets MDD and PSD were mutually exclusive.

MDD – consisted of Ç1 cough events from Ɒ1 patients. Cough events from this dataset were used to design the optimal model.

PSD – consisted of Ç2 cough events from Ɒ2 patients. Cough events from this dataset were used to test the designed model. Cough events from PSD were blind to the model design process.

Two expert scorers with 15 to 20 years of experience in paediatric respiratory diseases then scored the cough events from the two datasets into two classes, wet or dry. Scorers were blinded to the subject's history and diagnosis. This manual classification is considered the reference standard against which the automatic classification results are compared.

4.2.2. Design of cough sound classifier

To design a system for automatic classification of cough sounds, cough events from MDD were used. Let MDD1 be the subset of MDD containing those cough events on which both scorers agreed on the class of cough sounds. Let Ç11 be the cough events in MDD1. The cough events in MDD1 were used to design the automatic classifier model. This is a three step process.

[Step 1] Cough event Feature matrix computation: In this step, feature vector containing 'F' mathematical features is computed from each of the Ç11 cough events and a cough event feature matrix 'M_{MDD1}' of size, Ç11 × F was formed. To compute 'F' features from a cough event, the following below steps were used:

(i) Let x denotes a discrete time sound signal from a cough event $\hat{s}_{cg}[n]$.

(ii) Segment x into 'ν' equal size non-overlapping sub-segments. Let x_k represent the k^{th} sub-segment of x, where k = 1, 2, 3.

(iii) Compute the following features for each sub-segment and form a feature vector containing F features: bispectrum score (BSS), non-Gaussianity score (NGS), formant frequencies (FF), pitch (P), log energy (LogE), zero crossing rate (ZCR), kurtosis (Kurt), and twelve

Mel-frequency cepstral coefficients (MFCCs). The details of these features are explained in Appendix A1.

(iv) Repeat steps (i) through (iii) for all Ç11 cough events and form a cough event feature matrix M_{MDD1} of size Ç11 × F.

[Step 2] Automatic classifier design: In this study, a Logistic-regression model (LRM) is used as the pattern classifier. LRM is a generalized linear model, which uses several independent predictors to estimate the probability of a categorical event (dependent variable). In this work, the dependent variable Y is assumed to be equal to "one" (Y=1) for wet cough and "zero" (Y=0) for dry cough. A model is derived using a regression function to estimate the probability Y=1 (i.e. that a cough event belongs to the category of 'wet cough') given the independent variables (i.e. F features) as follows:

$$\text{Prob}\left(Y = 1 \mid_{f_1, f_2, f_3, \ldots, f_F}\right) = \frac{e^z}{e^z + 1} \tag{4-1}$$

$$z = \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \ldots + \beta_b f_F \tag{4-2}$$

In (4-1) and (4-2) f_1, f_2,…, f_F are the elements of feature vector (independent variables), β_0 is called the intercept and β_1, β_2 and so on are called the regression coefficients of independent variables. The Receiver-Operating Curve (ROC) analysis was used to select the optimal decision threshold θ from Y (that the cough is wet if Y is above θ, otherwise dry).

The data in matrix M_{MDD1} (Ç11 observations from F independent variables) was used and the leave-1-out cross validation (LOV) technique was adopted for LRM design. As the name suggests, the LOV technique involves using data from all cough events except one to train the model and one cough event to validate the model. This process was systematically repeated Ç11 times such that each cough event in MDD1 was used as the validation data once. This resulted in $Ł_{Ç11}$ number of LRMs.

To evaluate the performance of the designed $Ł_{Ç11}$, performance measures such as: sensitivity; specificity; accuracy; positive predicted value (PPV); negative predicted value (NPV); and Cohen's Kappa (κ) statistic were computed. The interpretation κ values are given in Appendix A2.

The design logistic regression model (LRM) is for:

(i). **Feature Selection**: Feature selection is a technique for selecting a sub-set of relevant features in order to build a robust learning model. Theoretically, optimal feature selection requires an exhaustive search of all possible subsets of features. However, to do so for a large number of features are computationally intensive and impractical. Therefore, a p-value was used to search for a satisfactory set of features. In LRM design, a p-value is

computed for each feature along with an indication of how significantly that feature contributed to the development of the model. Important features have low p-value. This property of LRM was used to select a reasonable combination of features (independent variables with low p-value) that facilitate the classification in the model during the training phase. The mean p-value is computed for 'F' features over Ç11 LRMs. The features are selected with a mean p-value of less than p_{ths}. Let F_B be the sub-set of the selected features from F.

(ii). **Robust LRM design**: A matrix is created that is M'_{MDD1} of size Ç11 × F_s from M_{MDD1}. Matrix M'_{MDD1} is a cough event feature matrix with only selected features F_B from Ç11 cough events in MDD1. Using M'_{MDD1} and adopting LOV, $Ç_{11}$ LRMs are retrained.

[Step 3] Selecting a good model from $Ł_{Ç11}$ LRMs: From $Ł_{Ç11}$ LRMs, one model was selected as the best using the K-mean clustering algorithm [124] to test on prospective study dataset PSD. In the K-mean clustering algorithm, the target is to divide *y* data points in *z*-dimensional space into *K* clusters, so that within the cluster sum of squared distance from the centroid is minimized.

The problem at hand was how to select a good model from the available $Ł_{Ç11}$ models. To do so, $Ł_{Ç11}$ models were divided in *d*-dimensional space into $K = 2$ clusters, that is, a high performance model cluster and a low-performance model cluster. A space dimension *d* was set equal to model parameters plus three performance measures (sensitivity, specificity and kappa). Then, from the cluster of the high performance models, the model that had the lowest mean square error value with respect to the centroid was selected. Let $\Re$ represent the selected LRM and $\theta_{\Re}$ is the corresponding probability decision threshold (the value is determined using ROC curves such that the classifier performance is maximized). Once $\Re$ is chosen, all the parameters of the model are fixed and use it for classifying cough sounds in the prospective dataset PSD.

4.2.3. Testing of selected LRM $\Re$

Following the procedure described in Section 4.2.2 [Step 1] and using the cough events from the dataset PSD, the cough event feature matrix M_{PSD} of size Ç2 × F is computed. Ç2 is the total cough events in PSD and 'U' is the number of feature vectors. M'_{PSD} from M_{PSD} is formed by selecting only robust F_B features. The selected LRM $\Re$ is used to classify data in M'_{PSD} into the classes of wet or dry. The decision process of the wet/dry class from the output of $\Re$ is as follows:

Let the output of the $\Re$ to a given cough input be $Y_{\Re}$. Then, the cough is classified as wet if $Y_{\Re} \geq \theta_{\Re}$ or dry otherwise.

The results of automatic classification by ℜ are compared with that of the expert scorers and the performance measures described in Section 4.2.2 [Step 2] are computed.

4.3. Results on automatic wet/dry cough classification

4.3.1. Cough sound datasets and agreement between expert scorers

This study uses sound recording data from Ɗ = 78 patients (41 males and 37 females). The mean age of the subjects was 2 years and 11 months. The age range of the subjects varied from 1 month to 15 years and they presented with diseases such as asthma, pneumonia, bronchitis and rhinopharyngitis. Table 4.1 gives the demographic and clinical details of the patients.

From Ɗ = 78 patients, a total of Ç = 536 cough events were analysed. On average, seven cough events per patients were analysed (minimum = 2 and maximum = 13). The dataset MDD has Ç1 = 385 cough events from Ɗ1 = 60 patients, while dataset PSD has Ç2 = 151 cough events from Ɗ2 = 18 patients.

Table 4.1: Demographic and clinical details of the subjects

GENDER	Male	41
	Female	37
AGE	Neonatal	2
	< 12 months	31
	< 60 months	29
	>= 60 months	16
DIAGNOSIS	Pneumonia	34
	Pneumonia + other	21
	Bronchitis	8
	Asthma	3
	Rhinopharyngitis	5
	Asthma + Rhinopharyngitis	1
	Others	6

Table 4.2 shows the contingency table between two scorers in classifying cough sounds from MDD and PSD, into the two classes of wet and dry. In dataset MDD, out of 385 cough events, scorers agreed Ç11 = 310 times (80.5%) on the classes of cough events which were used to form the

subset MDD1. In dataset PSD, they agreed 117 times out of 151 (77.5%). The Kappa agreement between Scorer 1 and Scorer 2 is 0.55 for MDD and 0.54 for PSD. Of the 310 cough events in MDD1, 82 belonged to the wet class and 228 belonged to the dry class. The MDD1 cough events were then used to design LRM models described in Section 4.2.2.

Table 4.2: Contingency table between human scorers for classifying coughs into wet/dry. $\kappa = 0.56$ and % agreement = 80.5% for MDD and $\kappa = 0.54$ and % agreement = 77.5 for PSD.

Dataset MDD					Dataset PSD				
		Scorer 1					Scorer 1		
		WET	DRY				WET	DRY	
Scorer 2	WET	82	55	**60%**	Scorer 2	WET	47	23	**67%**
	DRY	20	228	**92%**		DRY	11	70	**86.4%**
		80.4%	**80.6%**	**310**			**81%**	**75.3%**	**117**

4.3.2. Cough sound characteristics in the databases

The mean duration of dry cough in MDD1 was 260±77 ms (computed using 228 dry coughs) and that of wet cough was 238±54 ms (computed using 82 wet coughs). Figure 4.2 shows a typical example of the dry cough waveform and wet cough waveform from two patients, ids #35 & #38 respectively. The cough sound waveforms were generally clean with high signal to noise ratio (SNR). The mean SNR for MDD1 was 15.2±5.5 dB (maximum = 28.65 dB and minimum = 2.9 dB); the mean SNR for PSD was 18.6±4.5 dB (maximum = 27.8 dB and minimum = 11.1 dB). Figure 4.3 shows the histogram of SNR for the cough sound in MDD1 and PSD.

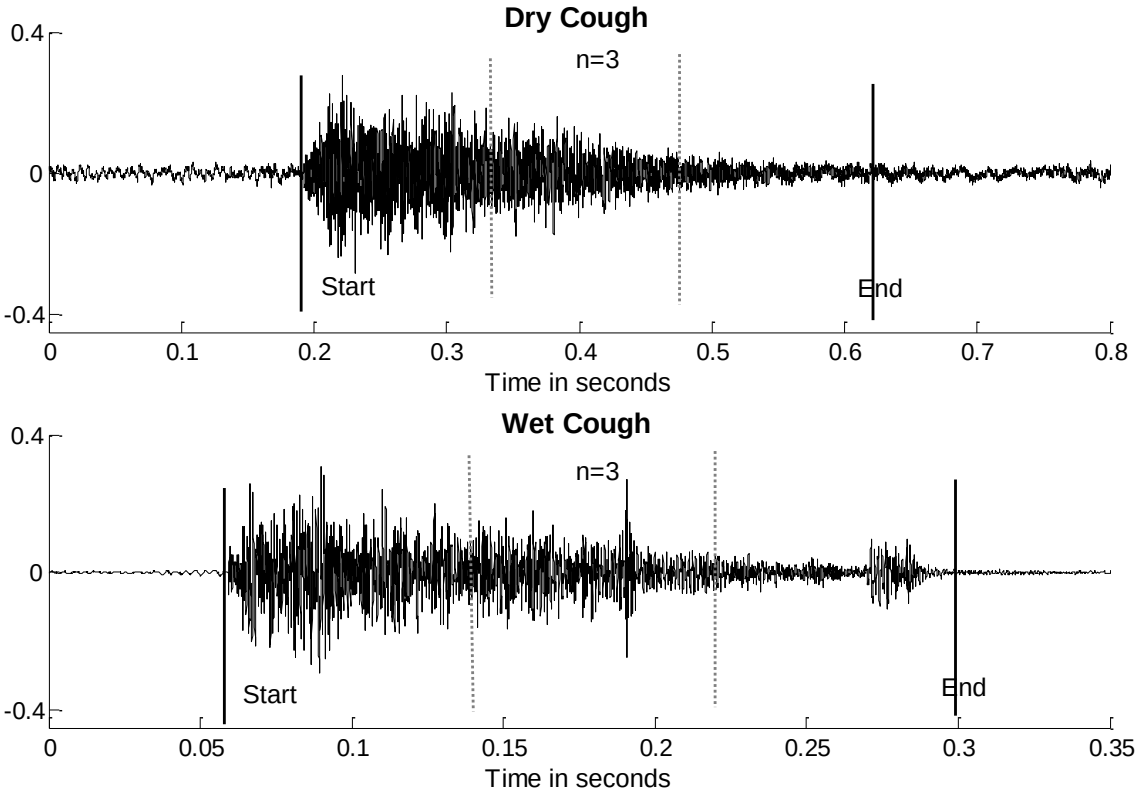

Figure 4.2: Typical examples of dry cough waveform and wet cough waveform from two patients, ids #35 & #38 respectively in the MDD dataset. Each cough segment is divided into ν (ν =3) sub-segments.

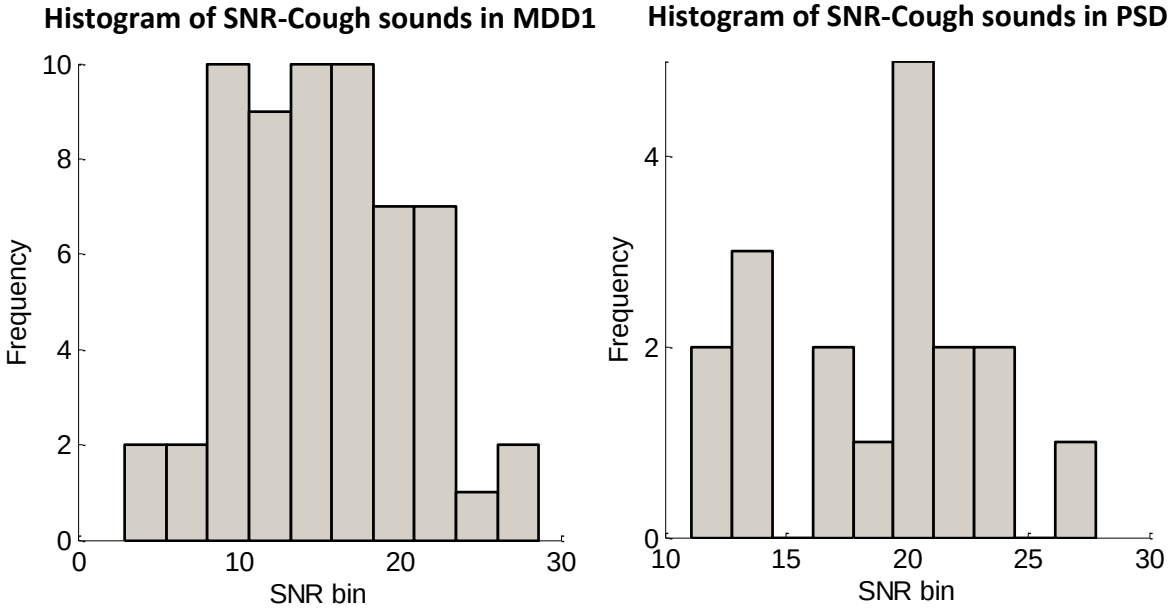

Figure 4.3: Histogram of signal to noise ratio (SNR) for the cough sound in MDD1 and PSD. The mean SNR for the cough sounds in MDD1 was 15.2±5.5 dB (maximum = 28.65 dB and minimum 2.9 dB) and that for PSD was 18.6±4.5 dB (maximum 27.8 dB and minimum = 11.1 dB).

Figure 4.2 exhibits the cough signals with the start and end markers. Following the method given in Section 4.2.2 [Step 1], the feature matrix M_{MDD1} was computed. The value $\nu = 3$ was used to divide each cough segment into three sub-segments. In the literature, clinicians and scientists alike have described cough sounds as consisting of three phases: an initial opening burst, followed by noisy airflow, and, finally, by glottal closure [125, 126]. It has been shown that these phases carry different significant information specific to the quality of cough, wet or dry. On this basis, each cough segment was divided into three sub-segments. Setting ν=3 led to a feature vector F of length 66 consisting of following features $(\nu \times 12\ MFCC) + (\nu \times 4\ FF) + ([\nu \times [BGS, NGS, P, LogE, ZCR,$ Kurt]). From Ç11 = 310 cough events and F = 66 features, cough event feature matrix M_{MDD1} was created.

4.3.3. Automatic classification using LRM

Feature Matrix and LRM performance during training stage: Following LOV technique, $Ł_{Ç11}$= 310 LRMs were designed. The mean training sensitivity and specificity for the 310 LRMs were 92±1% and 93±0.5% respectively. Validation sensitivity and specificity for these models were 62% and 84% respectively. Table 4.3–(A) gives the detailed classification results when all the F=66 features were used to train the LRMs.

Following the process described in Section 4.2.2 [Step 2] and using $p_{ths} = 0.06$, $p_s = 31$ features was selected. Figure 4.4 shows the mean p-value associated with F=66 features computed over Ç11=310 LRMs. All the features which have mean p-values of less than $p_{ths} = 0.06$ were selected. The selected features were one each from the Bispectrum score, kurtosis, and number of zero-crossings, two each from the non-Gaussianity score and log-energy, five from formant frequencies, and 19 from Mel-frequency cepstral coefficients. Table 4.4 gives details of the feature selected for designing the final LRM. According to this table, MFCC based features were the most dominant. Out of 31 selected features, 19 features were contributed from different MFCCs components. After MFCCs, formant frequencies made the second most dominant contribution with five features. Moreover, except for fourth formant frequency and pitch based features, which were completely omitted, all other features contributed with features from at least one sub-segment towards the building of the final LRM model.

Table 4.3: LRM performances before and after the feature selection. Statistics provided in the table are mean ± standard deviation. A 95% confidence interval for the mean of the training dataset is provided at bottom. For Scorer 1 and Scorer 2, the sample size is Ç1 = 385 cough events from Ð1 = 60 patients in dataset MDD. Out of 385 cough events, scorers had a wet/dry consensus on Ç11=310 cough events.

		Sensitivity	**Specificity**	**Accuracy**	**PPV**	**NPV**	**K**
(A) When all the features were used to develop LRM							
Consensus of scorer 1 & scorer 2	**Training**	91.76±0.68 [91.69-91.84]	92.65±0.45 [92.6-92.7]	92.45±0.5 [92.36-92.47]	81.80±1 [81.68-81.91]	96.90±0.3 [96.87-96.93]	0.8125±0.1 [0.8112-0.8138]
	Validation	62	84	78	59	86	0.46
Scorer 1 wet/dry class	**Training**	87.15±0.95 [86.9-87.4]	87.49±0.89 [87.26-87.72]	87.40±0.90 [87.17-87.63]	71.53±1.84 [71-72]	94.97±0.40 [94.87-95.07]	0.6977±0.02 [0.69-0.70]
	Validation	53	78	71	47	82	0.3
Scorer 2 wet/dry class	**Training**	81.96±1.01 [81.7-82.23]	82.24±0.97 [81.98-82.49]	82.14±0.98 [81.89-82.4]	71.83±1.37 [71.48-72.19]	89.18±0.78 [88.98-89.38]	0.6224±0.01 [0.6173-0.6276]
	Validation	45	67	59	43	69	0.12
(B) When selected all the features were used to develop LRM							
Consensus of scorer 1 & scorer 2	**Training**	87.36±0.61 [87.29-87.43]	87.82±0.43 [87.77-87.87]	87.70±0.46 [87.65-87.75]	72.07±0.87 [71.98-72.17]	95.07±0.25 [95.05-95.10]	0.7041±0.01 [0.7029-0.7053]
	Validation	81	83	82	63	92	0.58
Scorer 1 wet/dry class	**Training**	82.75±0.57 [82.60-82.89]	83.06±0.52 [82.92-83.19]	82.98±0.52 [82.84-83.11]	63.78±1.18 [63.47-64.08]	93.03±0.27 [92.96-93.10]	0.60±0.01 [0.59-0.60]
	Validation	76	79	78	57	90	0.5
Scorer 2 wet/dry class	**Training**	75.66±0.57 [7.5.51-75.81]	75.92±0.58 [75.77-76.07]	75.83±0.57 [75.68-75.98]	63.44±0.96 [63.19-63.69]	84.95±0.61 [84.79-85.11]	0.49±0.01 [0.4916-0.4975]
	Validation	72	73	72	59	82	0.43

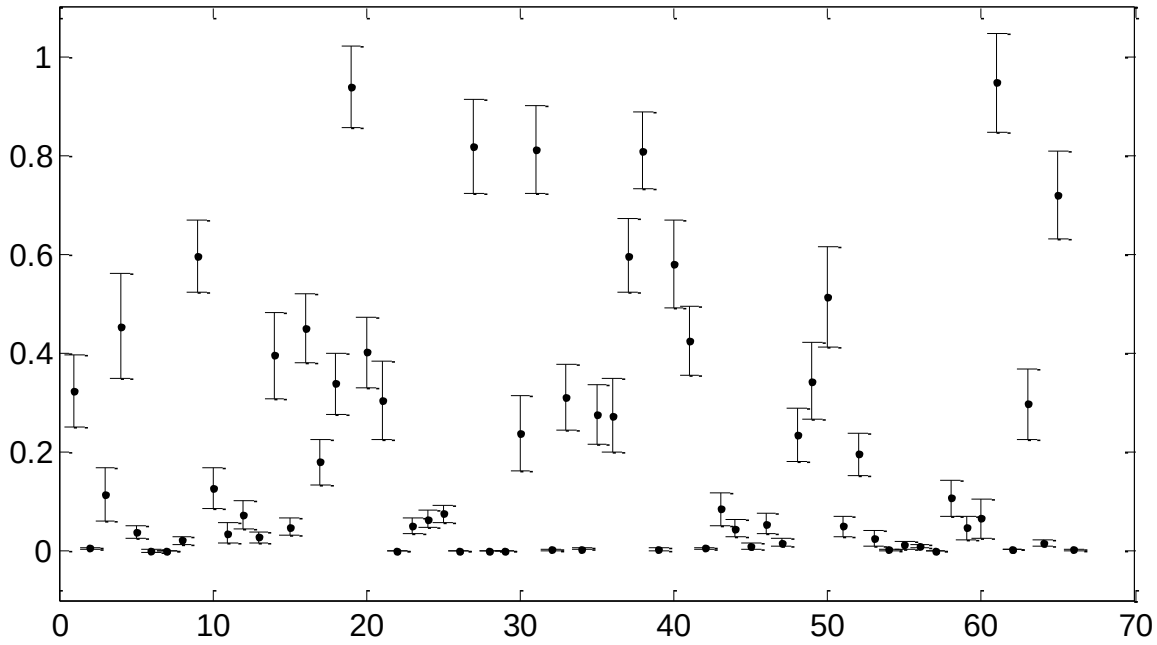

Figure 4.4: Mean p-value and standard deviation error bar, associated with F=66 features computer over 310 trained LRMs. The p-value indicates the associated significance level of a feature in the developing model.

When only selected features F_B were used to re-train LRMs, mean training sensitivity and specificity were recorded as 87±1% and 88±% respectively, while validation sensitivity and specificity were 81% and 83%. The validation kappa agreement between the LRM and scorers was 0.46 when all the features were used to train LRM; this increased to 0.58 when only selected features were used. Table 4.3(B) gives the detailed training and validation results after feature selection.

Selection of LRM (ℜ): From $Ł_{Ç11}$ = 310 designed LRMs using data from MDD1, the optimal model ℜ was selected using the K-mean clustering method as discussed in Section 4.2.2 [Step 3]. Models were clustered into two groups, a high performance model and a low performance model based on model parameters and performance measures. Of 310 models, 202 were clustered in the high performance model group, while 108 were in the low performance model group. LRM model #26 has the lowest mean square error value with respect to the centroid of the high performance models. This model ℜ was chosen and all its parameters were fixed for future use. ℜ was tested on the prospective dataset PSD.

Performance of ℜ on the prospective dataset PSD: Table 4.5 gives the classification results of ℜ against expert scorers. For Scorer 1, the wet/dry classification was used as the reference standard, and ℜ had the sensitivity of 77.5%, the specificity of 76%, and a kappa agreement of 0.47. For

Scorer 2, the results were sensitivity 75%, specificity 64%, and kappa 0.31%. When model ℜ was tested on only those events, in which Scorer 1 and Scorer 2 agreed on classification (117 cough events), sensitivity jumped to 84% and the kappa value went up to 0.51. Table 4.6 shows the contingency table.

Table 4.4: F = 66 features were computed from each cough segment by using ν = 3 at Section 4.2.2 [Step 1]. '√' indicates that feature was selected for designing the final model at Section 4.2.2 [Step 2].

Features	BSG			NGS			FF1			FF2			FF3			FF4		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Selected		√			√	√	√	√			√		√		√			
Features	Pitch			LogE			Kurt			ZCR			MFCC0			MFCC1		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Selected				√	√			√		√				√		√		
Features	MFCC2			MFCC3			MFCC4			MFCC5			MFCC6			MFCC7		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Selected			√	√		√		√	√	√	√				√		√	√
Features	MFCC8			MFCC9			MFCC10			MFCC11								
	1	2	3	1	2	3	1	2	3	1	2	3						
Selected	√	√	√		√			√		√		√						

LRM results when matched for Age and Gender: Table 4.7 shows the performance of the LRM on MDD1 and PSD when matched for age and gender. Due to the limited availability of data, only four divisions were considered: (i) male with age <= 60 months; (ii) female with age <= 60 months; (iii) male with age > 60 months; and (iv) female with age > 60 months. According to this table during the model designing stage, no significant difference was seen in the model validation performance across four divisions, as compared to when no division was considered, as seen in Tables 4.3 and 4.7(A). Similarly with the prospective dataset PSD, the selected model ℜ performed well across all divisions (Tables 4.5 and 4.7(B)), except in the third division (males with age > 60) where performance was very poor.

Table 4.5: Performance of ℜ on dataset PSD (prospective study dataset).

ℜ against individual scorer when tested on all the cough events (151) from DS2						
	Sensitivity	Specificity	Accuracy	PPV	NPV	κ
Scorer 1	77.5%	76%	76%	54%	90%	0.47
Scorer 2	75%	64%	67%	43%	87%	0.31
ℜ tested on only those events when both Scorer 1 and Scorer 2 agreed on class						
	84%	76%	78%	55%	93%	0.51

Table 4.6: Contingency table for selected LRM tested on dataset PSD. κ = 0.51. W = Wet, nW = not Wet.

		Scorers		
		W	nW	
LRM	W	26	21	55%
	nW	5	65	93%
		84%	76%	78%

Table 4.7: LRM validation results for the dataset MDD1 and the prospective dataset PSD with age and gender matched.

(A)						
Validation results for dataset MDD1. All the features were used to train the LRM						
	Sensitivity	Specificity	Accuracy	PPV	NPV	κ
Age <=60 months, Male (#121 cough events)	59%	83%	76%	57%	84%	0.41
Age <=60 months, Female (#145 cough events)	58%	88%	80%	63%	85%	0.47
Age >60 months, Male (#20 cough events)	89%	64%	75%	67%	87.5%	0.51
Age >60 months, Female (#24 cough events)	100%	83%	83%	20%	100%	0.28
Validation results for dataset MDD1. Selected features were used to train the LRM						
	Sensitivity	Specificity	Accuracy	PPV	NPV	κ
Age <=60 months, Male (#121 cough events)	73.5%	78%	77%	57%	88%	0.47
Age <=60 months, Female (#145 cough events)	84%	87%	86%	70%	94%	0.67
Age >60 months, Male (#20 cough events)	89%	64%	75%	67%	87.5%	0.51
Age >60 months, Female (#24 cough events)	100%	91%	92%	33%	100%	0.47
(B)						
Prospective Study dataset PSD						
	Sensitivity	Specificity	Accuracy	PPV	NPV	κ
Age <=60 months, Male (#36 cough events)	92%	87.5%	89%	78.5%	95%	0.76
Age <=60 months, Female (#27 cough events)	87.5%	95%	92.5%	87.5%	95%	0.82
Age >60 months, Male (#30 cough events)	50%	54%	53%	14%	87.5%	0.02
Age >60 months, Female (#24 cough events)	86%	71%	75%	54.5%	92%	0.48

4.4. Discussion

This study proposes an automated method to classify cough sounds into wet and dry categories. As far as I know, this is the first attempt to develop objective technology for the dry/wet classification of paediatric cough sounds. Besides being the first research of its type, it is also unique in that we proposed and validated methods to classify a given cough event into dry/wet groups; this is in contrast to previous research, which are limited to qualitatively describing characteristics of cough events pre-classified by a human observer. The results presented in this study are based on 536 cough events from 78 subjects, which is a considerably larger sample than that of existing research [121, 122] which use no more than 30 coughs in their descriptive analyses. For these reasons, the results cannot be directly compared against other work.

The reference method used for the assessment of my method is the subjective classification of cough sounds into wet/dry classes by two paediatric respiratory paediatricians from different countries. These scorers were blinded to the actual clinical diagnosis of the subjects. In an event by event cough classification, the two experts agreed with each other at a Moderate Level (kappa value of $\kappa = 0.54$). In [49], the inter clinician agreement for wet/dry cough is reported as $\kappa = 0.88$. However, it should be noted that, the clinicians assessed the wetness of cough at the patient level, but not at the individual cough level. When the agreement was computed between scorers at the patient level, the kappa value increased to $\kappa = 0.66$ (substantial agreement). These numbers further illustrate the subjective nature of the wet/dry classification.

The classifier technology was trained on coughs from the training set (dataset MDD) using only events where both scorers reached consensus. At the output of the training process, a good Logistic Regression Model ($\Re$) was identified and its parameters were fixed. The model was then tested on the Prospective Set (dataset PSD) in several different ways. The highest sensitivity and specificity (84% and 76%, respectively) of classification were achieved when $\Re$ was tested against consensus events within PSD. It is interesting to note that these numbers were consistently higher than what we got by testing against individual classification outcomes of each scorer.

Another salient feature of the method is that it has a high negative predictive value (NPV = 93%), when scorer consensus data is used as the ground truth. This means that if the model classifies a cough as dry, it is most likely that the two expert scorers would reach the same conclusion independently. However, the positive predictive value of the method compared to human scorers is lower (PPV = 55%). Thus, a sizable fraction of coughs classified by the model as wet ends up being consensus-classified as dry by human scorers. This phenomenon appears to be explained by the results presented by Chang et al. [49] which found that expert human scorers

underscore wet coughs. In [49] they systematically compared subjective wet/dry classifications of expert clinicians with bronchoscopy indications of airway mucus. They reported that clinicians' classification of dry cough do not necessarily indicate the absence of secretions. Certain situations in airways, such as when there are small amounts of secretions, may not be reflected in cough sounds at a sufficient magnitude to be detected by a human observer. One possible reason for a lower PPV value found in the method is this weakness in the gold standard, human scorers, used to generate the performance statistics. This hypothesis needs to be carefully validated against bronchoscopic findings in the future.

The ability to correctly detect airway mucus is particularly important in the management of suppurative lung diseases [49, 127]. Cough is an early symptom of diseases such as pneumonia, bronchitis, and bronchiolitis. The accurate assessment of this symptom is a crucial factor in the diagnosis of acute diseases and in the monitoring of chronic symptoms and treatment efficacy. It is known that, in children, wet coughs are more likely to be associated with lower respiratory tract infections [127]. The subjective classification of wet cough has low sensitivity as a method of detecting airway mucus, even in the hands of expert clinicians. Accurate, objective technology for the classification of dry/wet coughs is currently unavailable either at the commercial or research levels. To the best of my knowledge, this work is the first attempt in the world to develop such technology.

4.5. Conclusion

The proposed method in this study can classify cough sounds into wet and dry classes with high accuracy and good agreement with the assessment of paediatricians. This is the first known method for wet/dry classification, presented with complete training and testing results on significantly large cough samples. It is also the first effort to automate the wet/dry classification in the paediatric population with range of respiratory infectious diseases. It carries the potential for development as a useful clinical tool for long term cough monitoring, and in the assessment of treatment efficacy or in characterizing lower respiratory tract infections. It will be essentially useful in clinical or research studies where temporal patterns of cough quality (wet/dry) on an hour to hour basis are needed.

The methods proposed in this study should be available for simultaneous implementation with other potential technologies, such as microwave imaging and ultrasound imaging, which may be capable of detecting consolidations and mucus in lungs.

Novelty and Impact:

- First method to facilitate automated wet/dry coughs classification.
- Included a large number of non-voluntary/spontaneous cough sound samples obtained from paediatric populations.
- Can be developed as a tool to study temporal patterns of wet cough in long-term recordings.

Chapter 5

Cough Sound Analysis: An Alternative to the Bronchodilator Test in Remote Areas

Separating paediatric asthma from paediatric pneumonia is one of the major issues in remote areas. These diseases have overlapping symptoms, but require drastically different treatments. Existing guidelines for pneumonia classification in resource poor regions call for the use of a test called the bronchodilator test to separate asthma from pneumonia. It is time consuming, inconvenient, and not readily available in remote regions. In this chapter, I propose a substitute for the bronchodilator test based on the quantitative analysis of cough.

5.1. Introduction

Pneumonia and asthma are highly prevalent in paediatric populations [53, 109, 128]. These diseases contribute to the high mortality and morbidity of children, especially those who live in low income countries [53, 129]. In the first level health facilities of these countries, even primitive imaging and laboratory testing facilities are extremely rare. Further, it is also difficult to find trained healthcare personnel with expert clinical skills. The management of these diseases in such regions is largely dependent on community health workers who visit remote communities. To assist the community health workers in the fields, The World Health Organization (WHO) and United Nation International Children's Funds (UNICEF) have developed basic guidelines known as Integrative Management of Childhood Illness (IMCI) [55, 130].

IMCI guidelines describe the symptoms and procedures to clinically diagnose diseases in the paediatric population including pneumonia and asthma. According to the guidelines, the clinical signs of cough and/or difficulty of breathing are screening criteria for pneumonia. Fast breathing, defined as a breathing rate higher than a particular threshold (50 breaths/min in children younger than 12 months and 40 breaths/min in children between 12 and 60 months of age), is a determinant of whether pneumonia exists [55]. The existence of lower chest in-drawing indicates severe pneumonia. The diagnosis of asthma is established by symptoms of wheezing (often with cough),

fast breathing, hyperventilation of the lungs, chest wall in-drawing, and prolonged expiration [130]. Asthma and pneumonia have overlapping symptoms but require drastically different treatments.

The implementations of IMCI guidelines are successful in sorting out pneumonia and asthma from other respiratory diseases [59-61]. However, the similarity of the symptoms between pneumonia and asthma makes it difficult to differentiate between these two diseases with the result that asthma patients are often over-diagnosed as having pneumonia and receive unnecessary antibiotic treatments.

To address this problem, the updated WHO guidelines [67] recommends a test known as the bronchodilators test to differentiate asthma from pneumonia, in a population of subjects whose diagnosis is already narrowed down to the category: "either pneumonia or asthma".

As specified in IMCI guidelines [67], the bronchodilator test is administered to children exhibiting wheeze and fast breathing and/or lower chest in-drawing. After the first clinical observation by the physician, these children are given a bronchodilator via the oro-nasal route to dilate their airways. After 15 minutes, the clinical examination is repeated. If the symptoms that led to the bronchodilator prescription have disappeared, the patient is classified asthmatic [131]. Otherwise, the classification is pneumonia and antibiotics are prescribed [131].

Unfortunately, the bronchodilator test is time consuming. Bronchodilators and their delivery systems such as inhalers or nebulizers are also expensive and rare in resource-limited settings. Further, extra efforts to sterilize the bronchodilator deliver systems are required to avoid the spread of infections. An alternative for bronchodilator test is urgently required.

This chapter proposes a pioneering class of technology addressing these challenges. The target is to develop a cough-based technology as an alternative to bronchodilator tests used in the differential diagnosis of asthma from pneumonia in remote settings. The proposed method centres on cough sounds analysis to separate asthma patients from pneumonia patients.

Cough is one of the main symptoms of pneumonia and asthma. It is well known that cough sounds carry information related to respiratory diseases and researchers have attempted to investigate the acoustic properties of asthma coughs. One such study reported that the expulsive phase of asthma coughs containing quasi-periodic signal with fundamental frequency between 3 and 6 kHz. In a different study, asthma coughs were identified as sounds characterised by long duration wheezing [114, 132]. Asthma coughs have relatively high frequency components compared to cough caused by, for example, chronic bronchitis and tracheobronchial collapse syndrome. A common thread of these studies is that the acoustics of asthma coughs are seen as distinct from coughs in other diseases.

I hypothesize that cough sounds carry vital information that can be used to separate asthma from pneumonia. Support for this hypothesis comes from the patho-physiology of pneumonia and asthma, the physics of cough sound generation, and prior explorations [115]. The study in [19] showed that wavelet coefficients of cough sounds can be used to differentiate healthy, asthmatic, and chronic obstruction pulmonary diseases (COPD) subjects. Our own feasibility studies [115] indicated that cough carries vital information that can be extracted by quantitative analysis and used to screen pneumonia in remote regions.

In this work, I propose a Hidden Markov Model (HMM) based classifier for pneumonia-asthma classification. The HMM is used to model the temporal acoustic characteristic of cough sounds. Each cough sound is assigned with log-likelihood values corresponding to the likelihood that it will belong to either the pneumonia or asthma class.

To the best of my knowledge, the proposal to replace the bronchodilator test using a cough-based analysis is the first of its kind in the world. The proposed method is fully automated and does not use sensors that require physical contact with patients, so there is no need for sterilisation. These benefits make it easy to use in resource-poor regions by minimally trained personnel. The outcome of this work has the potential to transform the way the pneumonia/asthma is managed in remote resource-limited settings of the world.

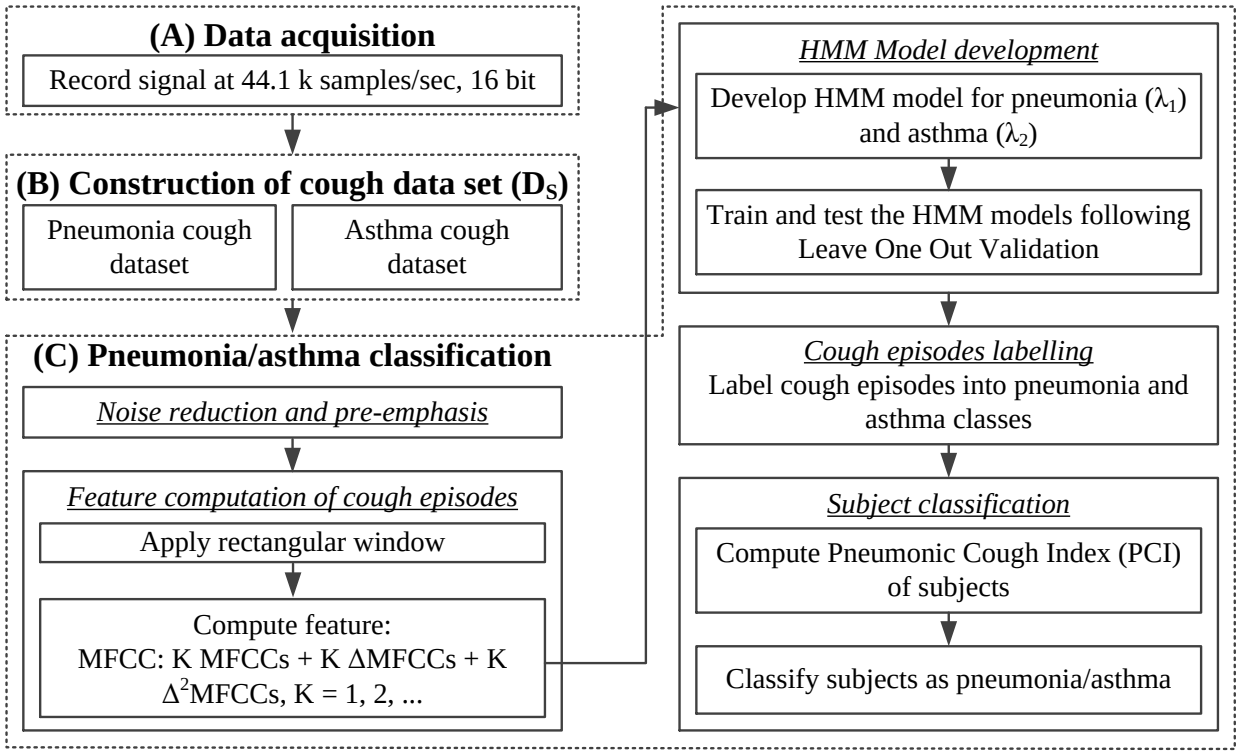

Figure 5.1: The block diagram of the proposed method. There are three main stages: (A) data acquisition; (B) construction of cough data set; and (C) pneumonia/asthma classification.

5.2. Material and method

The method proposed in this study is summarized in the overall block diagram in Figure 5.1. Each block represents a stage in the method: (A) data acquisition; (B) construction of the cough dataset; and (C) pneumonia and asthma classification. Details of these processes are described in Sections 6.2.1 to 6.2.3.

5.2.1. Data acquisition

The data for this work was obtained using the data acquisition systems and protocols described in Chapter 3.

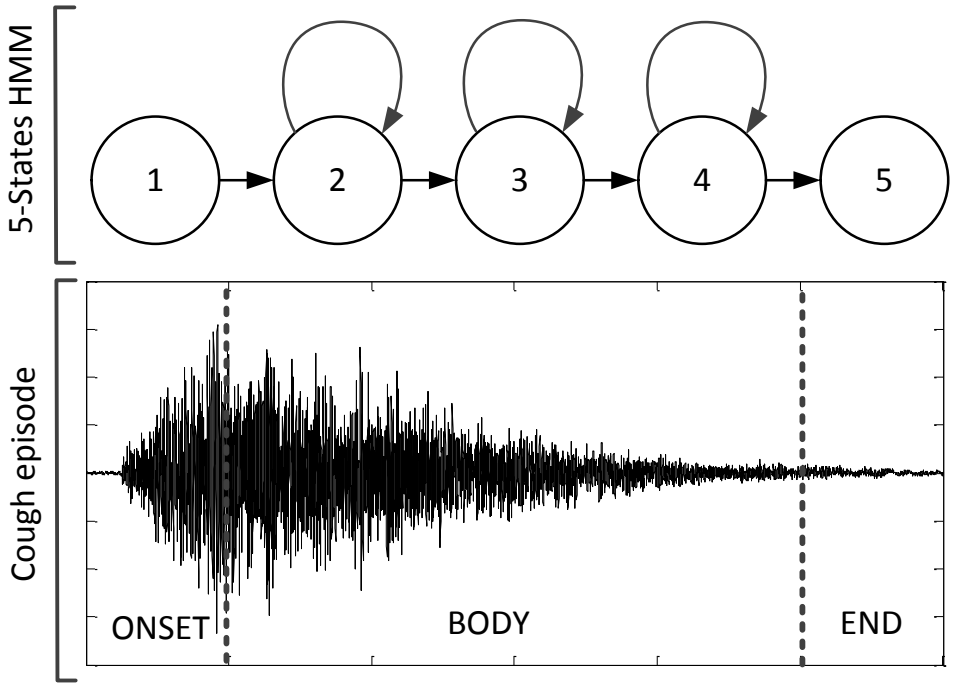

Figure 5.2: Five states of HMM are used to model onset, body, and end of a cough episode. States 1 and 5 are non-emitting states, while states 2, 3, and 4 are emitting states.

5.2.2. Construction of cough dataset

In this study, I involved Ɲ paediatric subjects (Ɲ = 20) admitted to hospital with respiratory complaints. The clinical diagnosis of pneumonia/asthma was established by professional paediatricians at Sardjito Hospital in Yogyakarta, Indonesia.

The cough dataset used in this study were constructed by manually picking Ç first cough episodes (Ç = 50) from each recording. The criteria of the cough selection are described in Section 3.3 of Chapter 3. If the number of coughs in a recording is less than ∂, then all coughs fulfilling the

criteria were used. Let DS represents the coughs obtained from this process with a size of Ç × Ð. These DS coughs were then used to train and test the Hidden Markov Model (HMM) based classifier for pneumonia-asthma classification. Details of this process are described in Section 5.2.3.

5.2.3. Design of pneumonia and asthma classifier

To classify pneumonia and asthma subjects I process the cough episodes through three steps (S1-S3) as follows:

(S1) Noise reduction and pre-emphasis:

The noise reduction process is described in Section 3.3 of Chapter 3. The noise reduction process implements a High Pass Filter and a Power Spectral Subtraction (PSS) filter. The filtered cough signal is denoted as $\hat{s}_{cg}[n]$. In this work, I further process the cough episodes through a first order pre-emphasis filter to enhance their high frequency components [133]. The output from the pre-emphasis process is denoted as $\dot{s}[n]$.

(S2) Feature computation of the cough episode:

The process of feature vector computation from $\dot{s}[n]$ is as follows:

(i). Apply a rectangular sliding window $w_r[n]$ of length N, generating data sub-blocks. Let $\dot{s}[n] = (|\dot{s}_1[n]|, \ldots, |\dot{s}_k[n]|, \ldots, |\dot{s}_K[n]|)$ represents the filtered cough sound where $\dot{s}_k[n]$ represents the k^{th} ($k = 1, 2, \ldots, K$) sub-block in $\dot{s}[n]$.

(ii). For each sub-block $\dot{s}_k[n]$, compute Mel-Frequency Cepstral Coefficients (MFCC). The MFCCs are spectral features computed from short-time sound signal. It approximates the auditory system behaviour by using non-linear frequency scale [133]. In this work, I used MFCCs to describe the temporal characteristic of cough episodes. Details of MFCC computation is given in Appendix A1.1.

Let Φ_k is the MFCC coefficients of k^{th} sub-block from a cough episode $\dot{s}_k[n]$. The MFCC features of a cough episode $\dot{s}[n]$ can be expressed in a matrix form as ($\Gamma = [\Phi_1, \Phi_2, \ldots, \Phi_k, \ldots, \Phi_K]$.

(S3) **HMM model development for classifying pneumonia/asthma**:

(i). **HMM structure**: In this study, a cough episode is represented as three stages, namely: Onset, Body, and End. To model the probable temporal characteristic of these stages, I developed i-states ($i = 5$) HMM. The illustration of the HMM model is shown in Figure 5.2. State 1 and 5 are non-emitting states, while States 2, 3, and 4 are emitting states that correspond to a particular respiratory tract configuration and model the probability characteristics of observation feature vectors Φ_k of each stage of cough. The Gaussian Mixture Model (GMM) is used to model the probability characteristics of the

observation MFCC feature vector Φ_k associated with each emitting state. The probability of Φ_k from state *i* of the HMM is governed by the output probability density function $b_i(\Phi_k)$ given in (5-1) which is defined as a multivariate GMM with *j* Gaussian mixture components (*j* = 1, …M, where M is the number of Gaussian mixture components, i.e. 1, 2, 4, and 6).

$$b_i(\Phi_k) = \sum_{j=1}^{M} m_{ij} b_{ij}(\Phi_k) \qquad (5\text{-}1)$$

where m_{ij} is the j^{th} mixture weight of the i^{th} state of HMM and $b_{ij}(\Phi_k)$ is defined as in (5-2).

$$b_{ij}(\Phi_k) = \frac{1}{(2\pi)^{\partial/2}|\Sigma ij|^{1/2}} \exp\left\{-\frac{1}{2}(\Phi_k - \mu_{ij})T(\Sigma i_j) - 1(\Phi_k - \mu_{ij})\right\} \qquad (5\text{-}2)$$

where μ_{ij} and Σ_{ij} are the mean and the covariance matrix of the j^{th} Gaussian distribution of the i^{th} state, and ∂ is the dimension of the feature vectors.

The HMM parameters (state transition and emission probability distributions) are estimated using the Baum-Welsh re-estimation procedure with Viterbi alignments.

Let $\Lambda = (\lambda_1, \lambda_2)$ denotes the set of respiratory diseases model of cough episodes where λ_1 and λ_2 respectively represent the HMM model for pneumonia and asthma. The likelihood of a cough episode with a sequence of *k* observation feature vector ($\Gamma = [\Phi_1, \Phi_2, \ldots, \Phi_k, \ldots, \Phi_K]$), from any given HMM model λ_y ($y = 1, 2$) is expressed in terms of the log-likelihood (*LL*) given in (5-3).

$$LL = \log P(\Gamma \mid \lambda_y) = \log P(\Phi_1, \Phi_2, ..., \Phi_K \mid \lambda_y) = \sum_{k=1}^{K} \log P(\Phi_k \mid \lambda_y) \qquad (5\text{-}3)$$

A Viterbi decoder return $\hat{\lambda}$ the best pneumonia/asthma model that give the highest log-likelihood score by computing $P(\Phi_k|\lambda)$ in terms of the HMM probability distributions according to (5-4).

$$\hat{\lambda} = \arg\max_{\lambda \in \Lambda} \log P(\Gamma \mid \lambda) \qquad (5\text{-}4)$$

(ii). **Pneumonia/asthma classification rule**: Let $\Omega = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_t, \ldots, \Gamma_T\}$ be a set of cough episodes from a patient, where Γ_t is the t^h cough episode and $|\Omega| = Ç$.

Rule 1: Cough episode classification

$$\forall \Re \in \Omega, if\ \lambda_y = \hat{\lambda} = \arg\max_{\lambda \in \Lambda} \log P(\Gamma \mid \lambda) then\ \Gamma \xrightarrow{\text{is attributed to}} y \qquad (5\text{-}5)$$

For all cough episodes Γ belong to set of Ω, if the cough model that produces the highest *LL* for cough episode Γ is the y^{th} cough model, then Rule 1 attributes Γ to the y^{th} cough model.

Rule 2: Pneumonia/asthma classification

Applying Rule 1 repeatedly on to Ω leads to a mutually exclusive and exhaustive partition of Ω into two model classes: pneumonia (λ_1) and asthma (λ_2). To classify a subject as pneumonia/asthma, I implemented a *Pneumonic Cough Index* (PCI). Let Q be the number of coughs classified as pneumonia and Ç be the total number of cough episodes analysed from a subject. The PCI can be computed as given in (5-6). Subjects that have a higher PCI than the optimum PCI threshold (γ) are classified as pneumonia and vice versa.

$$PCI = Q/Ç \quad (5\text{-}6)$$

In the classification process, I followed the 'leave one out' validation method where all subjects were used in training except one for testing. This process was systematically repeated such that each subject was used as the validation data once.

5.3. Results

5.3.1. Dataset

In this study, I used recordings from Ɒ = 20 subjects consisting of 8 males and 12 females. The age of the subjects ranges from 1 to 86 months (average age 25 months). In the dataset, the ratio of pneumonia and asthma subjects was equal (10 of each disease). Of 10 pneumonia subjects, chest x-ray was used to confirm nine of them.

According to the physical examination findings, nine of the pneumonia subjects and five of the asthma subjects had respiratory rates above the WHO threshold (respiratory rate greater than 50 breaths per minute in children younger than 12 months and greater than 40 breaths per minute in children aged 12 to 60 months). Fever (body temperature > 37.5°C) was presence in seven pneumonia subjects and five asthma subjects. Abnormal lung sounds found by clinicians in these subjects through a conventional stethoscope were crackles (pneumonia = 9 and asthma = 1) and wheeze (pneumonia = 2 and asthma = 7). Respiratory distress symptoms (sub-costal retractions) were present in all pneumonia subjects and in four asthma subjects. Three asthma subjects had

oxygen saturation levels below normal (< 96), while it was below normal value in only one pneumonia subject.

The total number of coughs in the data set DS was 739, with 462 coughs from pneumonia subjects and 277 coughs from asthma subjects. Over a 60 minutes period, the average number of coughs in pneumonia subjects was larger than that in asthma subjects (46 coughs and 28 coughs, respectively).

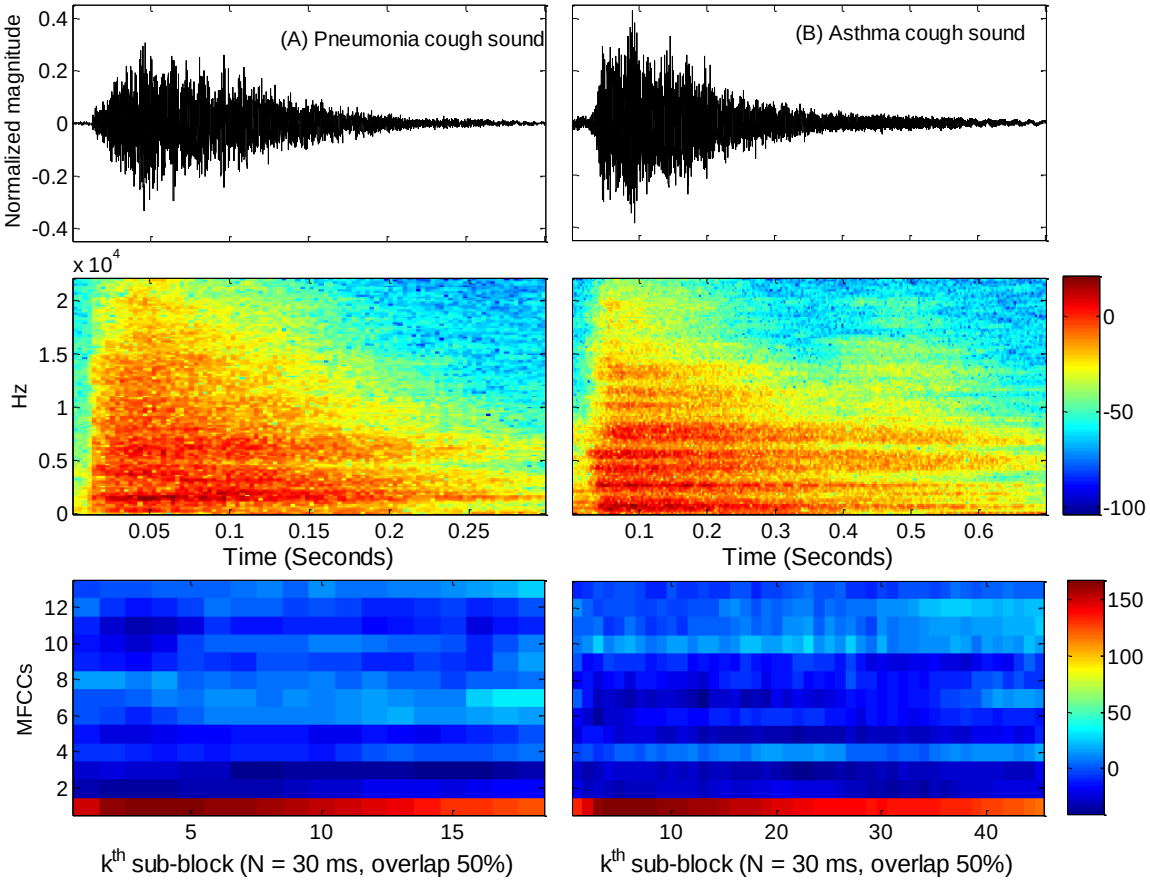

Figure 5.3: Illustration of signals, spectrogram, and MFCCs of coughs from pneumonia and asthma subjects.

5.3.2. Characteristic of cough in pneumonia and asthma

In step (S2)-(i) of Section 5.2.3 I used a sliding rectangular window $w_r[n]$ of length N to segment the data in each cough episode $\hat{s}[n]$ into sub-blocks. Then I computed the features for each sub-block. The performance of the method in different length of N is presented in Section 5.3.3.

Figure 5.3 shows the signals, the spectrograms, and the MFCCs of coughs from pneumonia and asthma subjects. From the spectrogram, one can see that the asthma cough has longer duration with

clear formant/harmonics from the beginning to the end of cough. In contrast, the energy of pneumonia cough beyond 10 kHz is higher compared to asthma cough.

The MFCCs in Figure 5.3 was computed using N = 30 ms (1232 samples). Pneumonia and asthma coughs show the high intensity in their first MFCC. However, pneumonia has a relatively high magnitude as well at the 6th, 7th and 8th of the MFCCs, while the asthma cough shows high magnitude at the 4th, 10th, 11th, and 12th of the MFCCs.

5.3.3. Pneumonia asthma classification

To investigate the optimum sub-block size, the size of the rectangular sliding window (N) was varied from 20 ms to 60 ms in steps of 10 ms. To compute the features from a cough episode, the rectangular window was shifted from the beginning to the end of the cough episode with 50 percent overlap. From each sub-block, a feature vector containing 39 features was computed. Thirty-nine features are 13 of MFCC and their first and second order differentials.

Table 5.1: The accuracy of cough episodes classification into pneumonia and asthma classes follows **Rule 1** (in %). The results are presented in mean ± standard deviation.

	Number of GMM	Sub-block sizes (milliseconds)				
		20	30	40	50	60
Pneumonia	1	74.3±21.2	82.7±12.6	82.5±14.3	80.1±15.6	80.4±12.9
	2	72.9±23.7	76.3±17.0	76.8±20.1	77.6±18.0	78.0±16.8
	4	69.8±27.1	73.1±22.5	74.1±23.3	71.5±22.9	75.9±22.1
	6	72.7±23.9	78.1±20.3	75.4±21.2	75.1±17.1	75.3±18.9
Asthma	1	52.6±21.2	48.4±12.6	44.5±14.3	47.3±15.6	49.5±12.8
	2	45.8±23.8	50.0±17.0	45.2±20.1	51.3±18.0	51.1±16.8
	4	42.7±27.1	48.0±22.5	47.1±23.3	51.1±22.9	42.5±22.1
	6	42.3±23.9	41.5±20.3	43.9±21.2	48.9±17.1	37.7±18.8

Table 5.1 shows the accuracy of cough episodes classification into pneumonia and asthma classes using four different numbers of Gaussian Mixture Model/GMM (1, 2, 4, and 6) following Rule 1. From the Table 5.1, it can be seen that in the HMM model for pneumonia (λ_1), the optimum accuracy was obtained at sub-block size 30 ms with single GMM (mean 82.7%, standard deviation

12.6%). The table also shows a decreasing trend in accuracy with the addition of GMM components (2, 4, and 6).

In the HMM model for asthma (λ_2), the optimum accuracy of cough episodes classification at sub-block size 20 ms with single GMM is 52.6 percent. However, the standard deviation of the accuracy in this set is very high (21.2%). The optimum accuracy with the smallest standard deviation can be found at sub-block sizes either 30 ms or 60 ms with one or two GMMs. In this set, the accuracies are slightly lower by 1.5 to 4.2 percent but the standard deviations are 4.2 to 8.6 percent lower. Similar to λ_1, increasing the number of GMM by 2 and 4 also reduced the accuracy of λ_2.

Following Rule 2 in Step 3 (S3) in Section 5.2.3, I computed the Pneumonic Cough Index (PCI) [113] and computed the performance of the method in classifying pneumonia/asthma subjects at sub-block sizes 30 ms and 60 ms with one and two GMMs. I varied the threshold and plotted the receiver operating characteristic (ROC) curve. Figure 5.4 shows the ROC curve of the testing set computed using the 'leave one out' validation procedure. As can be seen, there are overlapping areas in the ROC curves at both sub-block sizes with one/two GMMs.

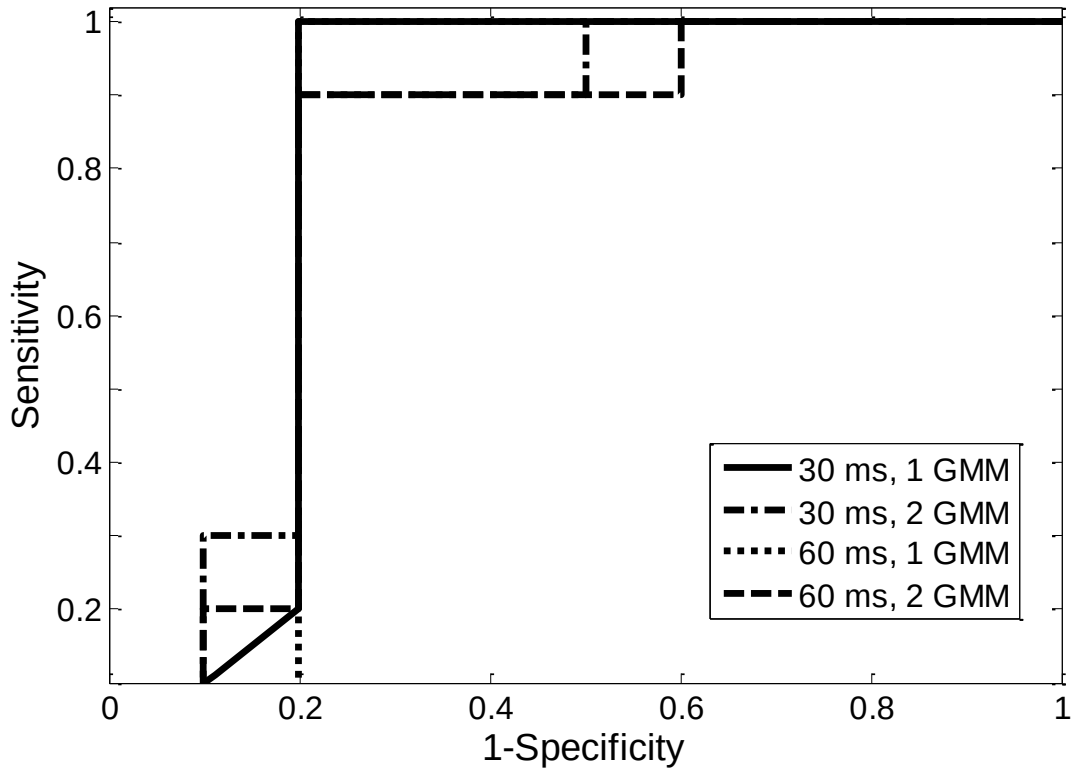

Figure 5.4: ROC curve of pneumonia and asthma classification computed using the Pneumonic Cough Index (PCI) defined as ratio of the number of pneumonia coughs divided by the total number of coughs in a patient.

Table 5.2: The performance of pneumonia and asthma classification computed using *Pneumonic Cough Index* (PCI) follows Rule 2. The γ, Sens, Spec, Acc, PPV, NPV, and κ, respectively represent the optimum PCI threshold, Sensitivity, Specificity, Accuracy, Positive Predictive Values, Negative Predictive Values, and Cohen's Kappa statistics.

Sub-block sizes	Number of GMM	γ	Sens	Spec	Acc	PPV	NPV	κ
30	1	0.60	100.0	80.0	90.0	83.3	100.0	0.80
	2	0.61	90.0	80.0	85.0	81.8	88.9	0.70
60	1	0.61	100.0	80.0	90.0	83.3	100.0	0.80
	2	0.66	90.0	80.0	85.0	81.8	88.9	0.70

Table 5.2 shows the performance of the pneumonia and asthma classification computed using PCI at sub-block sizes 30 ms and 60 ms. At the defined optimum PCI thresholds (γ), in both sub-block sizes with one GMM, our method achieved sensitivity and specificity of 100 and 80 percent, respectively. The increment of GMM to 2 decreases the sensitivity by 10 percent. Table 5.2 also shows that the agreement between our results and the diagnosis from the hospital is high as shown by Kappa statistics 0.7 to 0.8.

5.4. Discussion

Although pneumonia is a major cause of mortality among children, it is often over-diagnosed at the cost of under-diagnosing asthma in resource poor regions of the world. The main reasons for this include: unavailability of advance diagnostic tools, difficulty in finding trained healthcare personnel with expert auscultation and clinical skills, and, most importantly for my research, the presence of overlapping clinical symptoms between the diseases.

In my own database, all the subjects presented with cough and/or difficult breathing, which are central symptoms for screening pneumonia according to the WHO algorithm used by community workers. More than 50 percent of asthma subjects had respiratory rates higher than the recommended WHO threshold for diagnosing pneumonia. Moreover, 30 percent of asthma subjects had sub-costal retraction, which is an indication of severe pneumonia and requires immediate administering of antibiotics according to the WHO algorithm. This shows that if the WHO guidelines were followed, at least 50 percent of asthma subjects would be misclassified as

pneumonia in my database. These results are in line with [61] which reported that out of 200 children diagnosed with pneumonia based on the WHO criteria, 46 percent actually had asthma.

One study [63] suggested adding fever to the WHO algorithm for pneumonia diagnosis to improve its specificity. However, 44.4 percent of asthma subjects also had fever, which makes fever an unreliable symptom to use in separating asthma from pneumonia. Our feasibility work on pneumonia community screening [115] also indicated that fever had limited use in detecting pneumonia with higher specificity.

At present to separate pneumonia cases from asthma cases, WHO recommends the bronchodilator test [67, 68] as an additional diagnostic criterion for asthma. However, bronchodilators and their delivery systems, such as inhalers or nebulizers, are costly. Further, the sterilization requirements for inhalers/nebulizers (to avoid the spread of infections) add extra costs to the bronchodilator tests. The bronchodilator test is not an efficient use of time either, as it can take up to 45 minutes to complete the test. More than 75 percent of deaths due to pneumonia and asthma occur in developing and low-income countries where first level medical services, if available, are sub-optimal at best. There is an urgent need for an alternative method to the bronchodilator test, which is not only cost effective, but also easily deployable in settings where resources are limited.

In this chapter, I discussed an innovative method of separating pneumonia cases from asthma using cough sounds only. The novelties of the proposed method lie in the classification technique as well as in the clinical applications. This is the first effort in this field to address fundamental problems with pneumonia/asthma classification due to the overlapping symptoms between these diseases.

My methods, when rigorously validated, are expected to provide an unorthodox new approach to detecting pneumonia patients in a mixed population of pneumonia and asthma subjects. In other words, the method may function as an alternative to an accurate, low-cost bronchodilator test that can be field-deployed in resource-poor remote regions of the world. My method uses non-contact measurements; therefore, it does not need extensive sterilization. Cough sound analysis can be fully automated and implemented as a self-standing package on a smart phone. This makes the technology easy to deliver and maintain in remote regions where it is really needed.

While I developed my technology largely targeting field-deployment in remote regions, its applications reach beyond that context. I believe it has applications in respiratory units of healthcare facilities and offices of primary care paediatricians, even in the developed world.

My pneumonia/asthma classification method based on cough analysis implements Hidden Markov Models (HMMs) incorporated with Mel-frequency cepstral coefficients (MFCCs). Cough

shares some similarities with speech. It consists of temporal structures and can be encoded as a sequence of spectral vectors spanning a specific frequency bandwidth. HMMs are capable of providing a natural framework for constructing such models [134].

As the feature set, MFCCs and its derivatives were implemented to describe the temporal patterns of cough. The MFCCs use the non-linear frequency scale to approximate the human auditory system [133]. By listening to the patient's cough sounds, trained clinicians and/or caretakers are able to differentiate coughs in several different categories such as wet and dry. In children, the characteristic of cough may suggest a specific aetiology [47, 110]. For example, wet cough may be correlated with protracted bacterial bronchitis/sinusitis. These facts show that the human auditory system can be used to support diagnosis. Therefore, in this work, we implemented MFCC to capture the features of cough sounds.

I tested my method on 20 subjects following the 'leave one out' validation method, and achieved high sensitivity, specificity, and Kappa. These results support the previous study [115] and indicate that cough sound analysis has the potential to be implemented as a substitute for the bronchodilator test to classify pneumonia and asthma.

5.5. Conclusion

In this study, I proposed a novel method for classifying pneumonia subjects and asthma subjects using cough sounds. This is the first effort in this field to address fundamental problems with classifying pneumonia and asthma due to the overlapping symptoms of these diseases. The results show that the proposed method has the potential to be implemented in resource-limited settings as a support for the existing guidelines from the WHO and as an alternative for the bronchodilator test.

Novelty and the impact:

- The first study to investigate cough sounds from paediatric pneumonia and asthma.
- Use of cough sound analysis to separate pneumonia from asthma
- Potential to be developed as a low cost system that can substitute for the bronchodilator test in remote resource-poor areas.

Chapter 6

Automatic Cough Segmentation in Paediatric Populations

Cough is the dominant single symptom leading to physician visits throughout the world. Cough carries vital information on the paediatric respiratory system, yet the quantitative analysis of cough is still in its infancy. At present, medical practitioners and researchers alike still rely on manual cough identification, which is laborious and time consuming. In this chapter, I develop an automated paediatric cough identification and segmentation technique based on Time-Delay Neural Networks (TDNN). The method can analyse a continuous sound recording and extract cough events while discarding other sounds, such as speech and ambient noise. The method has the potential to serve as a real-time paediatric cough-counting device and as the front end of a system to diagnose diseases such as pneumonia and asthma.

6.1. Introduction

Cough quantity, defined as the frequency of cough events over a given time interval, is one of the important markers evaluated by paediatricians during consultation sessions. This information can be used to determine the nature (e.g., acute, chronic) and the severity of coughs, as well as to monitor the efficacy of a treatment [20]. However, to obtain this information, paediatricians rely heavily on subjective reports provided by patients or their carers. There is a great need for an automated device capable of counting the number of coughs, especially for childhood disease. More importantly, technology capable of automatically extracting cough events from paediatric recordings is urgently needed in order to facilitate the diagnosis of diseases such as pneumonia.

Several approaches have been taken to develop automated cough counting systems (e.g. Hull Automatic Cough Counter (HACC), Leicester Cough Monitor (LCM), LifeShirt, VitaloJAK, and PulmoTrack). The performances of these devices are varied. The HACC claimed 80 percent sensitivity and 96 percent specificity [32]. The figures for LifeShirt, PulmoTrack, LCM, and VitaloJAK are (78%, 99%), (94%, 96%), (85.7%, 99.9%), and (97.5%, 97.7%) respectively [33-36]. To extract the sound events, HACC computed the standard deviation of cough sound intensities

[32], while VitaloJAK [135] and LCM [136] used the sound intensities themselves. This makes these methods susceptible to variations in recording conditions, such as the distance between the microphone and the patient, the sensitivity of the recording instruments, and the sound level of the coughs being recorded. Both LifeShirt [137] and PulmoTrack [138] described cough as having a distinguishable humped structure in the waveform. Then they identified coughs by counting the number and/or measuring the slope of these humps. None of these commercial devices has been tested on paediatric populations or on subjects with diseases such as pneumonia. In paediatric subjects, the cough sound intensities and waveform-shapes may vary widely. Therefore, intensity or simple waveform-shape based methods are unlikely to be optimal and are less likely to be robust in field use.

Cough recording on children, especially the younger ones, pose several additional challenges. Younger children are unable to produce voluntary coughs upon request. Any method targeting paediatric populations should be capable of using spontaneous coughs recorded over a period of time. In paediatric recordings, crying, vocalization, and grunting are found abundantly, and are intermixed with cough sounds. Consequently, technologies developed for adults are unlikely to be optimal for use on children.

Existing commercial cough counting devices such as VitaloJAK [135], LifeShirt [137], and PulmoTrack [139] employ contact sensors. While the use of contact sensors may have some advantages, they also carry several drawbacks. Contact sensors, compared to non-contact (free-air) microphones are robust against background sound propagated through air. However, they are more vulnerable to sound conducted through tissue and bones; spurious rubbing sounds due to sensor movement can also be an issue. In infectious diseases, elaborate efforts are needed to avoid cross contamination of patients through contact instrumentation. Furthermore, in paediatric subjects, contact sensors can also be difficult to attach because of patient discomfort.

I address these issues and propose a novel technology for the automated segmentation of cough events from recordings obtained using non-contact microphones in a paediatric ward. In particular:

- I design algorithms to target the paediatric population (age < 6 years), addressing a fundamental gap in current technology.
- I develop a method for the segmentation of cough events, with algorithms capable of discounting background sounds, such as crying, vocalization, and grunting.
- I develop new techniques that are robust against the variation of cough sound intensity levels and waveform-shapes. This is an unprecedented approach in this field, to the best of my knowledge.

The method has the potential to be developed as an automated cough counting device as well as the front-end of a cough based diagnostic system.

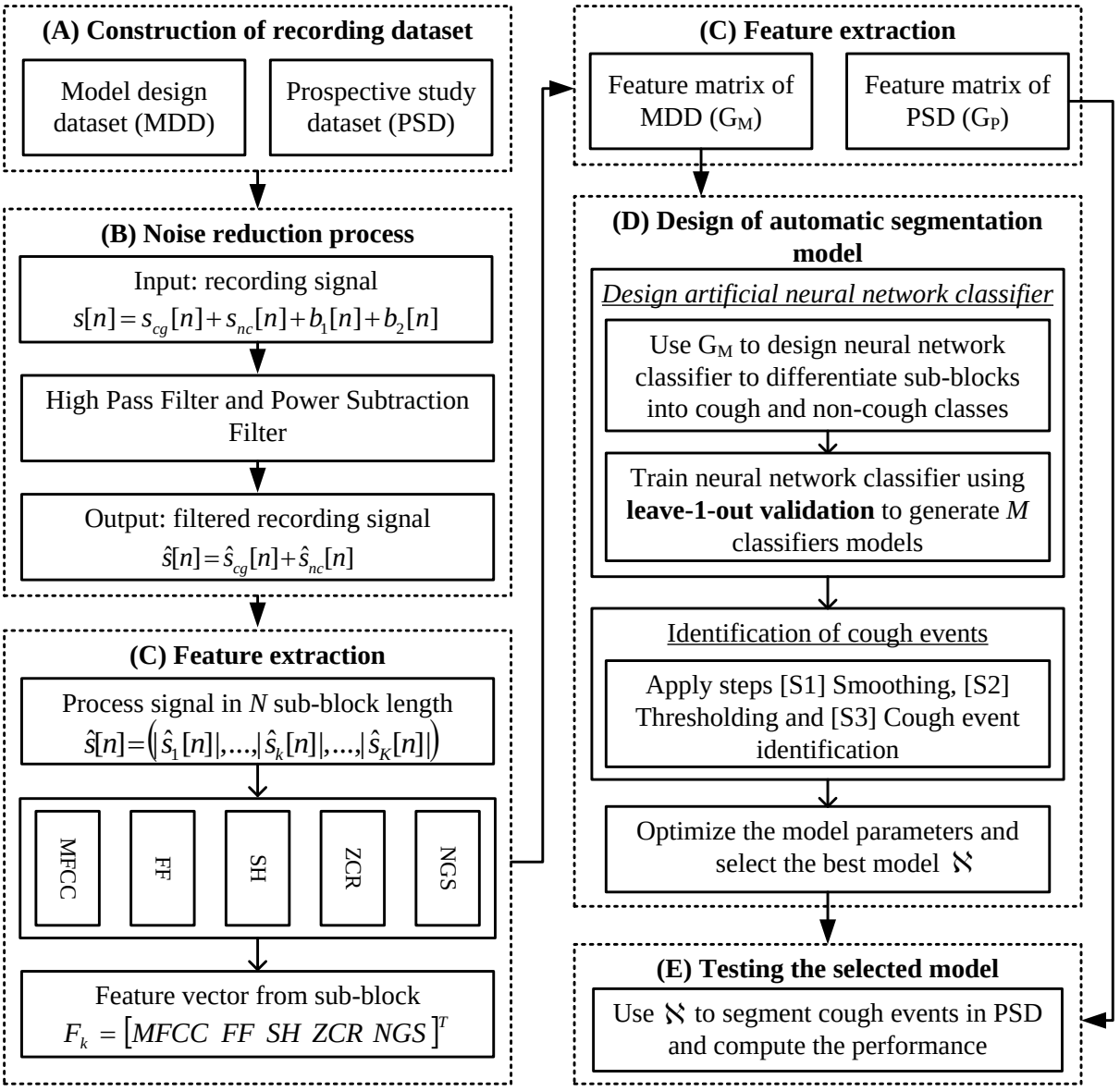

Figure 6.1: The block diagram of the automatic segmentation proposed algorithm.

6.2. Material and method

The block diagram of the overall method is shown in Fig 6.1. The method is comprised of five main processes: construction of the recording dataset; noise reduction; design of feature extraction from sub-blocks of data; design of the automatic segmentation model; and testing of the best selected model. Details of the process are described in Sections 6.2.1 through 6.2.5.

6.2.1. Construction of recording dataset

The recording dataset for this study was obtained by following the procedure described in Chapter 3. The dataset was divided into two categories, namely:

(i). **Model Design Dataset (MDD)**: Dataset MDD denotes the training dataset that was used to train the neural network (NN) classifier. The main criterion in designing MDD is that it should contain the whole range of cough types and their variations, as well as non-cough sounds expected in the practical setting. That way, the NN can learn the characteristics of the variety of coughs and learn to differentiate them from non-cough sounds. Dataset MDD was designed by manually picking representative cough and non-cough sounds from each subject in the training dataset. All such sound events were then concatenated as a single data stream to form the set MDD. Thus, the dataset MDD is not a natural sequence of sound events, but the combination of a large number of handpicked cough and non-cough sounds. To develop MDD, I took cough and non-cough sound samples from $\eth 1 = 10$ subjects. The overall length of MDD was 15 minutes.

(ii). **Prospective study dataset (PSD)**: Dataset PSD denotes the testing dataset and contains the actual sound stream recorded in the hospital. This way, the results on PSD can be taken as a true indication of performance in the clinical environment. The testing set PSD included the first 60 minutes of the sound streams recorded from $\eth 2 = 14$ subjects in the testing set. The total duration of PSD was 840 minutes.

Subjects in dataset MDD and PSD were mutually exclusive. The division of the subjects in two datasets was based on the order in which they presented to the respiratory clinic of the hospital.

In both MDD and PSD, cough events were manually identified and marked. To define the beginning and end of cough segments, the scorer carefully listened to the sounds and simultaneously looked at the time domain waveform displayed on computer screen. This manual identification of cough events is used as the gold standard against which the results of the automatic classification were compared.

The recordings in MDD and PSD were then passed through noise reduction process. The process is described in the following section.

6.2.2. Noise reduction

To reduce the noise, the recording signal $s[n]$ is processed through a High Pass Filter and Power Spectral Subtraction Filter as described in Section 3.4 of Chapter 3.

Let $\hat{s}[n]$ be the estimate of the recording $s[n]$ after the high pass and power spectral subtraction filters. It consisted of $\hat{s}_{cg}[n]$ and $\hat{s}_{nc}[n]$ to represent the estimates of cough sounds and non-cough sounds respectively. The signal of interest is the cough sound estimate $\hat{s}_{cg}[n]$. In this chapter, an automated method is proposed to extract cough sounds $\hat{s}_{cg}[n]$ from the recording $\hat{s}[n]$. To do so, the sound features of $\hat{s}[n]$ is first computed and then processed through a classifier. The next section describes the feature extraction method used in this work.

6.2.3. Feature extraction

To obtain the features of the sound signal, I applied a rectangular sliding window $w_r[n]$ of length N (N = 882 samples, equal to 20 ms) to $\hat{s}[n]$, generating data sub-blocks. Let $\hat{s}[n] = (|\hat{s}_1[n]|, \ldots, |\hat{s}_k[n]|, \ldots, |\hat{s}_K[n]|)$ represents the filtered sound recording where $\hat{s}_k[n]$ represents the k^{th} (k = 1, 2, …, K) sub-block in $\hat{s}[n]$. For each sub-block $\hat{s}_k[n]$ I computed the following features: Mel-frequency cepstral coefficients (*MFCC*), formant frequency (*FF*), zero crossing rates (*ZCR*), non-Gaussianity scores (*NGS*), and Shannon entropy (*SH*), and then formed a feature vector $F_k = [MFCC\ FF\ ZCR\ NGS\ SH]^T$. Details of these features are presented in Appendix A.1.

In this study, I extract such feature vectors from the Model Development Dataset (MDD) and the Prospective Validation Set (PSD). Let the feature set extracted from the k-th sub-block of MDD be denoted by $F_{k,M}$ and that from the PSD be denoted $F_{k,P}$. Next, I form an overall feature matrix for $\hat{s}[n]$, based on the feature vectors of sub-blocks as: $G_M = (F_{1,M}, F_{2,M}, \ldots, F_{k,M}, \ldots, F_{K,M})$ and $G_P = (F_{1,P}, F_{2,P}, \ldots, F_{k,P}, \ldots, F_{K,P})$, where G_M and G_P represent feature matrices for MDD and PSD.

The cough feature matrix estimated from the set MDD, G_M, was then used to train the automatic classifier model to classify sound data in a sub-block $\hat{s}_k$[n] into the classes of cough (CG) and non-cough (NC). Details of this process are described in Section 6.2.4. The matrix G_P from the set PSD was used for the prospective testing of the trained models.

6.2.4. Design of automatic cough segmentation model

The automatic cough segmentation method is a two stage process: classification of sound features into cough (CG) and non-cough (NC) classes, and identification of cough events. The description of these processes is given in Section 6.2.4(A) and 6.2.4 (B).

(A). Design of neural network model to classify a sub-block $\hat{s}_k[n]$ into the cough (CG) and non-cough (NC) classes

In this study, I investigate the use of an Artificial Neural Network (ANN) as the CG/NC classifier at the sub-block level. I used the ANN inspired by the capability of human brain to

recognize different types of cough sounds, regardless of their intensity, duration, or wetness. Moreover, ANN has the advantage of classifying data using non-linear decision boundaries, based on a process of supervised learning with a set of given examples. In this work, I used the particular form of an ANN known as a Time Delay Neural Network (TDNN) [140] that has been used in speech recognition applications. TDNN is capable of classifying sub-blocks $\hat{s}_k[n]$, discounting temporal translations of the input feature set [140].

The TDNN structure is comprised of an input layer (L_i), two hidden layers (L_{h1} and L_{h2}), and an output layer (L_o). The number of neurons in L_i, L_{h1}, L_{h2}, and L_o are 110, 20, 10, and 1 respectively. I used a linear activation function for neurons in the Lo layer and sigmoid activation functions for neurons in the L_{h1} and L_{h2} layers. To determine initial weights and bias values, I used the Nguyen-Widrow initialization method [141]. For updating weights during the training process, I employed the resilient back propagation (RPROP) algorithm [142].

I trained the TDNN to classify each sub-block in $\hat{s}[n]$ into CG/NC class. Let $Q = [F_{k-2}\, F_{k-1}\, F_k\, F_{k+1}\, F_{k+2}]$ represents the feature vectors of sub-block $\hat{s}_{k-2}[n]$, $\hat{s}_{k-1}[n]$, $\hat{s}_k[n]$, $\hat{s}_{k+1}[n]$, $\hat{s}_{k+2}[n]$, respectively. I used Q as input to the TDNN for classifying the k^{th} sub-block $\hat{s}_k[n]$ into CG/NC class. This process was repeated for $k = 3, 4, \ldots, K\text{-}2$ to cover the whole signal $\hat{s}[n]$ represented by $\hat{s}[n] = (|\hat{s}_1[n]|, \ldots, |\hat{s}_k[n]|, \ldots, |\hat{s}_K[n]|)$. An illustration of this process is shown in Figure 6.2.

In training and optimizing the TDNN, I used the matrix G_M and adopted the 'leave one out' validation technique. This involves using feature matrices from all the subjects in MDD except one to train the TDNN model, and validate the model using the remaining subject. This process was systematically repeated M times (M = 10), such that each subject in MDD was used as the validation data once. This resulted in M neural network models.

Let $u_v[k]$, $v = 1, 2, \ldots, 10$ represents the output of the v^{th} TDNN model. In the following section, I describe the method to identify the beginning and the end of a cough events using $u_v[k]$.

(B). Identification of the beginning and the end of coughs events

The identification of the beginning and the end of cough events from the output of TDNN models $u_m[k]$ was carried by following steps (S1) through (S3):

(S1). Smoothing process - Pass $u_m[k]$ through a moving average filter with the tap length β. Let the output of this filter be $\tilde{u}_m[k]$.

(S2). Thresholding - Apply a threshold value ρ to $\tilde{u}_m[k]$. The sub-blocks with $\tilde{u}_m[k]$ above the threshold ρ were assigned the decision $Y = 1$, otherwise $Y = 0$. The sub-blocks with value $Y = 1$ are candidate cough event sub-blocks.

(S3). Cough event identification – A cough event is identified if q number of consecutive sub-blocks are candidate sub-blocks (i.e. $Y = 1$) and following conditions are satisfied:

Condition 1: $\tau_{min} \leq \tau_q \leq \tau_{max}$, where τ is the total time duration of q consecutive sub-blocks with $Y = 1$. τ_{min} and τ_{max} respectively are the minimum and maximum cough sound durations computed from the cough events in MDD.

Condition 2: $\delta_q > \delta$, where δ_q is the root mean square value of q consecutive sub-blocks in $\tilde{u}_m[k]$ with $\gamma = 1$ and δ is the threshold root mean square value.

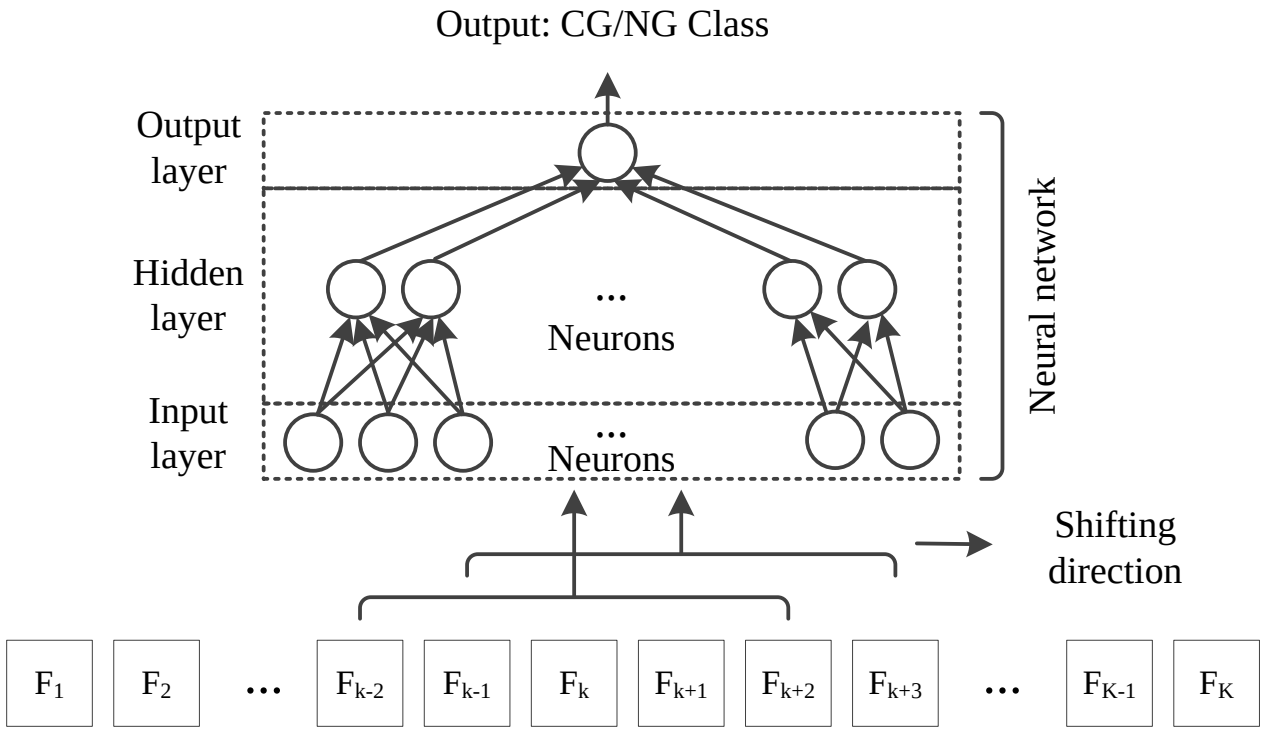

Figure 6.2: The structure of TDNN. It is comprised of an input layer, a hidden layer, and an output layer. Five successive inputs are used as input to TDNN to determine the class of a sub-block. The process starts from the beginning of recordings and shifts to the end of recording.

All the parameters including β, ρ and δ were optimized using data in MDD for maximizing the classifier performance. The steps for optimizing these parameters are as follows:

(P1). Optimizing β - Set parameter ρ and δ at certain values then vary β from 1 to 30. Select the optimum β values at the maximum performance in the validation set.

(P2). Optimizing ρ - Use the optimized β and δ values from step (P1), and then vary λ from 0.001 to 1. Select the optimum λ at the maximum performance in the validation set.

(P3). Optimizing δ – Use the optimized β and ρ values from steps (P1) and (P2), and then vary δ from 0.001 to 1. Select the optimum δ at the maximum performance in the validation set.

To evaluate the performance of the designed 10 TDNN models and proposed method for cough event identification, performance measures such as sensitivity, specificity, accuracy, and Cohen's Kappa statistic were computed. The performances were computed at the sub-block level as well as the cough event level by comparing the output of the algorithm with the reference scoring from the human observer.

From the 10 TDNN models, I selected the model that gave us the best performance as obtained in the 'leave-one-out' validation technique on MDD. Let ℵ represents the selected TDNN model with β_s, ρ_s, δ_s as its corresponding parameters. In Section 6.2.5, I describe the performance of ℵ in segmenting coughs from the prospective study data set (PSD).

6.2.5. Testing the selected model ℵ on PSD

The process of cough segmentation in PSD is as follows. Let G_P be a feature matrix computed using sound data from a subject in dataset PSD, following the process described in Section 6.2.4. Apply the model ℵ to G_P and automatically classify the cough sounds into classes CG/NC at the sub-block level; identify cough events following the steps given in Section 6.2.4. Repeat this for all the patients (ℶ2 = 14) in PSD. Compare the results of automatic segmentation with that of manual segmentation and evaluate the performance, at sub-block as well as event levels.

6.3. Results

6.3.1. Dataset

In this work, I included cough sound data from ℶ = 24 subjects (13 males and 11 females). The age of the subjects spanned from 3 months to 71 months (14 subjects < 12 months and 10 subjects > 12 months). After the clinical evaluation and laboratory tests, 18 were diagnosed with pneumonia, and one each with rhinopharyngitis, nasopharyngitis, tonsillopharyngitis, pulmonary hypertension, bronchiectasis, and bronchitis.

The training dataset MDD (ℶ1 = 14 subjects) has 656 cough and 2297 non-coughs events spread over a 15 minute period. Each subject contributed, on average, 66 coughs to the set MDD. The minimum length of cough event (τ_{min}) was 180 ms and the maximum (τ_{max}) was 720 ms.

The prospective dataset PSD has total sound data of 840 minutes duration from Ω2 = 14 subjects. In this dataset, there were 1434 coughs and 40144 non-cough sounds. The coughs have minimum and maximum length of 160 ms and 800 ms, respectively. On the average, there were 96 coughs in one hour of sound data from each subject.

The length distribution of coughs is illustrated in Figure 6.3.

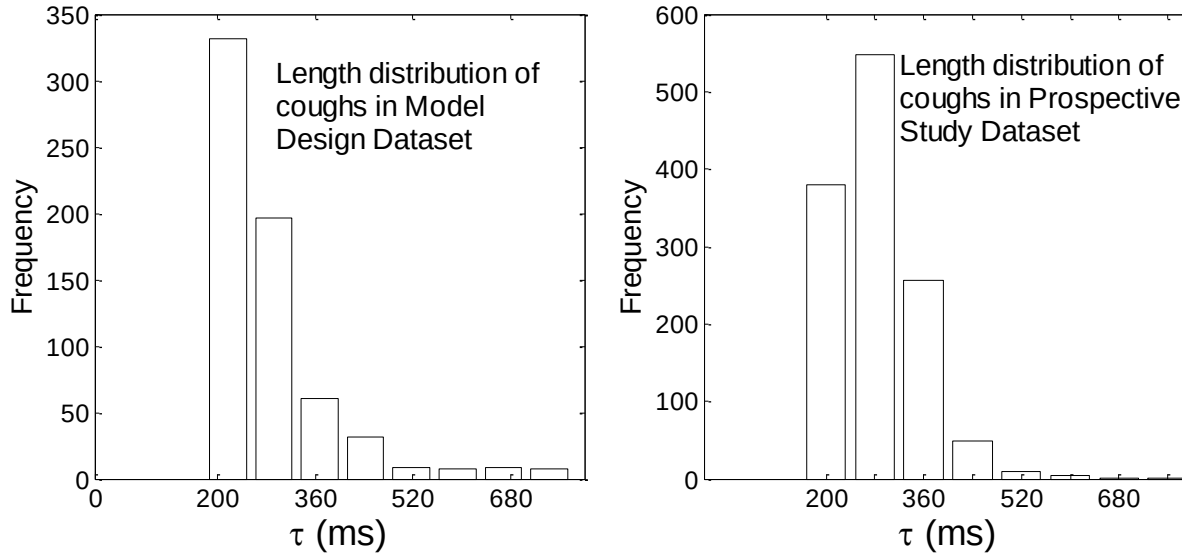

Figure 6.3: Histogram of the cough duration in MDD (a) and PSD (b). The minimum (τ_{min}), maximum (τ_{max}), and mean length of cough were 180 ms, 720 ms and 283 ms, respectively, in MDD and 160 ms, 800 ms and 291 ms, respectively, in PSD.

6.3.2. Parameter optimization of the classifier model using the set MDD

Following the process described in Section 6.2.2 and using the data from MDD, a cough feature matrix G_M was created. G_M was then used to train the automatic classifier model (Time-Delay Neural Network, TDNN) to classify sound data in sub-blocks $\hat{s}_k[n]$ into two classes: cough (CG) and non-cough (NC).

I varied the length of sub-blocks (N) to optimize the classification performance (sensitivity, specificity) of 10 TDNN models. The results are shown in Tables 6.1 and 6.2. Table 6.1 shows that the sub-block size N ranging from 20 to 40 ms gave a similar performance. Therefore, I set sub-block size = 20 ms in Section 5.2.2B. The reason behind this choice is that a shorter sub-block can give a better time resolution. Table 6.2 shows TDNN model performance when different combinations of feature sets were used in the training. From Table 6.2 it can be seen that when all features were used to train TDNN, it gave the best performance, both in training and in 'leave one

out' validation. For this reason, results reported in this study use TDNN models trained with all these features.

Table 6.1: Performances of TDNN in training dataset MDD using four different sub-block sizes. Statistics provided in the table are mean ± standard deviation. Tr, Va, Sens, and Spec respectively indicate training set, validation set, sensitivity, and specificity.

		Sub-block size			
		10ms	20ms	40ms	60ms
Tr	Sens	90.56±1.79	92.92±2.13	92.92±1.94	74.72±2.55
	Spec	90.57±1.79	92.92±2.13	92.93±1.94	74.72±2.55
Va	Sens	84.09±3.39	87.16±3.11	87.84±2.71	67.49±3.27
	Spec	84.11±3.39	87.17±3.11	87.91±2.69	67.56±3.27

Table 6.2: Performances of TDNN on training dataset MDD, using different combination of features. Statistics provided in the table are mean ± standard deviation. *FF* = Formant frequency, *SH* = Shannon entropy, *ZCR* = Zero Crossing Rate, *NGS* = Non-Gaussianity Score, and *MFCC* = MFCCs. Tr, Va, Sens, and Spec, respectively indicate training set, validation set, sensitivity, and specificity.

		Features				
		FF	*FF, SH*	*FF, SH, ZCR*	*FF, SH, ZCR, NGS*	*FF, SH, ZCR, NGS, MFCC*
Tr	Sens	71.68±3.12	73.29±3.09	79.52±2.72	81.09±2.38	92.92±2.13
	Spec	71.68±3.12	73.29±3.09	79.52±2.72	81.10±2.38	92.92±2.13
Va	Sens	68.47±4.43	70.21±5.20	77.32±5.34	78.89±4.26	87.16±3.11
	Spec	68.48±4.43	70.24±5.21	77.35±5.34	78.92±4.25	87.17±3.11

Parameters β, ρ and δ were optimized to achieve the best classification performance following the process described in Section 2.4.4. Figure 6.4 shows parameter optimization results for β, ρ and δ. According to Fig 6.4(A), at low β values, the classification sensitivity is high. It starts decreasing as β increases above 10. No significant variation in validation specificity can be seen with β value.

In case of the parameter ρ (Fig 6.4(B)), the validation sensitivity increases with ρ reaching the peak value at ρ = 0.116 and then starts decreasing. On the other hand, the validation specificity shows a consistent increase with an increase in ρ. For the parameter δ, validation sensitivity shows no variation until δ = 0.229 and then starts decreasing. Contrary to this, validation specificity starts increasing after δ = 0.229. Using these curves we set β_s = 9, ρ_s = 0.116 and δ_s = 0.327.

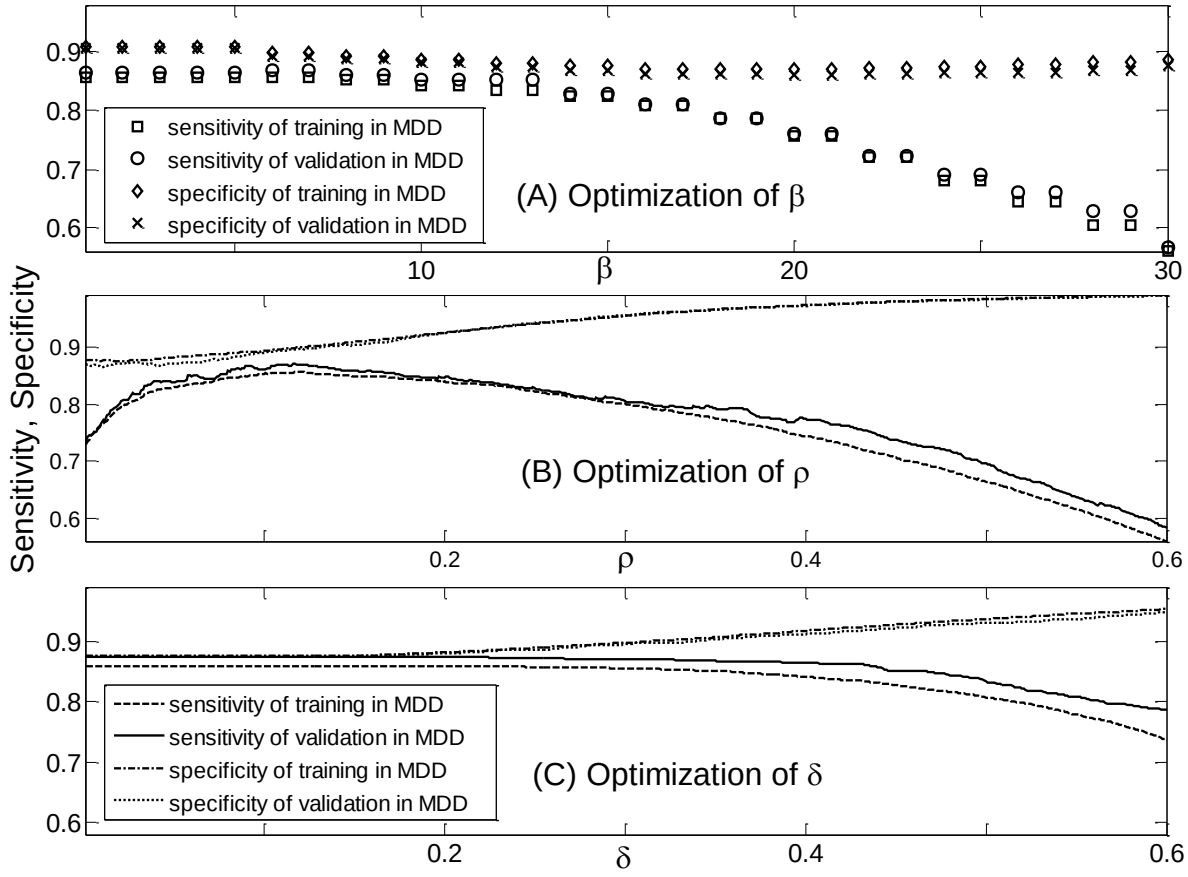

Figure 6.4: Illustration of classifier parameters (β, ρ, δ) optimization in model design dataset (MDD).

6.3.3. Segmentation results on training/validation dataset MDD

In this section, I illustrate the cough segmentation process and present the segmentation results on dataset MDD. The segmentation results are presented at the sub-block level, as well as the cough event level.

I illustrate a typical cough sound segment from MDD and the output of automatic segmentation at different stages of processing in Figure 6.5. Figure 6.5(A) shows the filtered signal $\acute{s}[n]$ while Figure 6.5(B) shows the output signal $u_m[k]$. According to Figure 6.5(B), $u_m[k]$ is high (very close to

1) during cough events (CG1, CG2, CG4, CG5, and CG6). Even though the intensity of CG5 is low (Figure 6.5(A)), $u_m[k]$ is still high. In Figure 6.5(A), CG3 is a cough event mixed with a speech signal and, in that case, $u_m[k]$ has a low value.

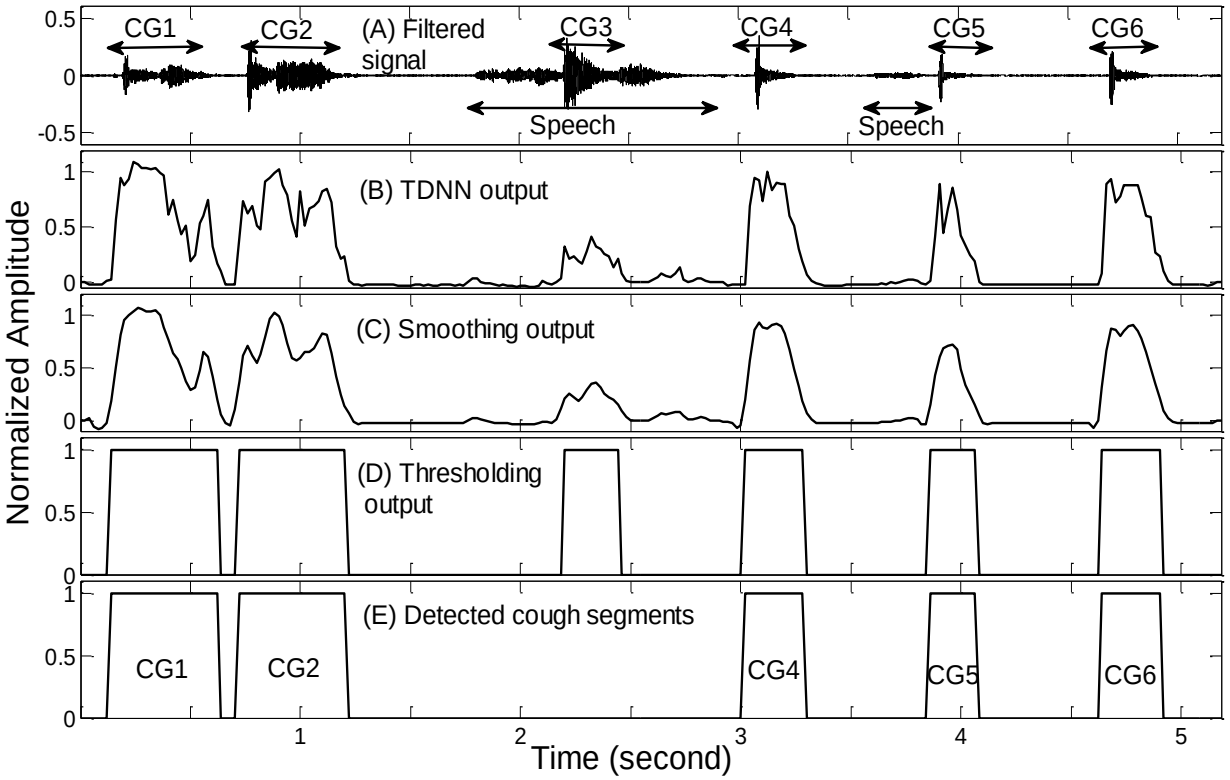

Figure 6.5: An illustration of cough segment identification from the model design dataset MDD. The cough sounds are indicated by CG1 through CG6. (A) output of noise reduction process, (B) output of TDNN, (C) output of moving average filter (step (S1) in Section 6.2.4, (D) output of thresholding (step (S2) in Section 6.2.4), (E) cough segments identified after step (S3) in Section 6.2.4.

To determine the beginning and the end of cough segments from $u_m[k]$, I followed the process described in Section 6.2.4(B) steps (S1) through (S3). Figures 6.5(C) through 6.5(E) shows the outputs of those steps. Figure 6.5(C) shows the output of the smoothing process $\tilde{u}_m[k]$. According to Figure 6.5(C), the smoothing process improves the shape of the signal by reducing the spurious transitions of $\tilde{u}_m[k]$, as in the case of CG5.

Figure 6.5(D) shows the candidate cough segments after applying the threshold. At step (S3) of Section 6.2.4(B), only the candidate segments fulfilling criteria ($\tau_{min} \leq \tau_q \leq \tau_{max}$ and $\delta_q > \delta$) were classified as cough segments. Figure 6.5(E), shows the output after step (S3). According to Figure 6.5(E), segments CG1, CG2, CG4, CG5, and CG6 were correctly identified as cough events. The

event CG3, which has speech segments, was not identified as a cough event. These results indicate that the proposed automated algorithm is capable of rejecting cough events corrupted with speech sounds and is robust against the variation of cough sound intensities.

Table 6.3: The performance of the algorithm on the training dataset (MDD) and prospective study dataset (PSD) using the optimized parameters (δ_s = 0.327, when β_s = 9, ρ_s = 0.116). Statistics provided in the table are mean ± standard deviation.

Training results for 10 TDNN models on dataset MDD				
(A) Performance computed at sub-block level				
	Accuracy	Sensitivity	Specificity	Kappa
Training	89.4±1.0	85.5±2.3	90.5±0.8	0.71±0.03
Validation	89.4±3.3	84.6±8.6	90.4±3.9	0.71±0.06
(B) Performance computed at event level (overlap > 50%)				
	Accuracy	Sensitivity	Specificity	Kappa
Training	93.6±1.2	91.3±1.9	94.8±1.2	0.82±0.03
Validation	93.9±2.0	89.8±7.0	94.8±1.7	0.82±0.06
Testing results using selected TDNN model M_s (δ_s, β_s, ρ_s) on dataset PSD				
(C) Performance computed at sub-block level				
	Accuracy	Sensitivity	Specificity	Kappa
Testing	99.1±0.5	83.7±8.6	99.3±0.4	0.60±0.1
(D) Performance computed at event level (overlap > 50%)				
	Accuracy	Sensitivity	Specificity	Kappa
Testing	97.4±1.1	92.8±6.8	97.5±1.1	0.65±0.1

In Tables 6.3(A) and 6.3(B), I present segmentation results obtained from the 10 TDNN models in the 'leave one out' validation. Table 6.3(A) shows the segmentation performance computed at the sub-block level. At the sub-block level, the algorithm achieved a sensitivity and specificity of 85.5±2.3 percent and 90.5±0.8 in the TDNN training process. The trained model resulted in a sensitivity and specificity of 84.6±8.6 and 90.4±3.9 percent respectively, in the 'leave one out' validation process. Kappa agreement between automated algorithm and manual scoring was 0.71±0.06 during the 'leave one out' validation.

Table 6.3(B) shows the performance of the automated algorithm in identifying cough events. For the work of this study, a true positive was detected if the event identified by the automatic classifier had at least 50 percent overlap with the manual scorer. The 'leave one out' validation sensitivity and specificity in identifying cough events were 89.8±7.0 and 94.8 percent respectively. The validation Kappa agreement between automated algorithm and scorer was 0.82±0.06 percent, which signifies a substantial agreement.

6.3.4. Segmentation results on prospective study dataset PSD

Following the process in Section 6.2.4, the TDNN model ℵ (with all the parameters fixed) was constructed and tested on the set PSD. Table 6.3(C) shows the performance of ℵ at the sub-block level. It achieved an accuracy, sensitivity, and specificity of 99.1±0.5, 83.7±8.6, and 99.3±0.4 percent, respectively. The Kappa agreement between ℵ and manual scoring was 0.61±0.13. Table 6.3(D) presents the performance of ℵ in identifying cough events in PSD. The automated algorithm achieved 97.4±1.1 percent accuracy, 92.8±6.8 percent sensitivity, 97.5±1.1 percent specificity, with a Kappa agreement of 0.65±0.1 in identifying cough events.

According to Table 6.3, the performance of the automated algorithm is better on PSD than on MDD. The major reason for this is the way the MDD and PSD have been designed, as described in Section 6.2.1. The MDD (training set) consists of manually picked sounds, with a target of presenting the TDNN with a range of sound episodes observable in a real recording. To make the learning process robust, in MDD I purposely included even difficult examples such as low SNR data and low intensity coughs. In the case of PSD, I used the actual recorded data streams, where such difficult events naturally occur at a much lower rate. Even though this strategy led to a slightly lower performance on the training set, it also led to a better generalization capability, which helped the M_s to produce a better outcome on PSD.

Table 6.4 presents the segmentation results in three age groups (age < 12 months, 12 to 36 months, and 36 to 72 months). From Table 4, it can be seen that the accuracy and specificity of the automated algorithm is similar across all age groups. The sensitivity in the age group 12 to 36 months is about 2 percent higher compared to the other age groups. The results show that my automated algorithm is not significantly affected by the age variations.

Table 6.4: Performance of the selected TDNN model M_s (δ_s, β_s, ρ_s) on the Prospective Study Dataset (PSD).

	Less than 12 months	12 to 36 months	36 to72 months
Accuracy	97.0±1.4	97.1±0.2	97.7±0.3
Sensitivity	91.4±8.1	94.6±4.7	92.2±7.8
Specificity	97.3±1.4	97.4±0.6	98.3±1.0

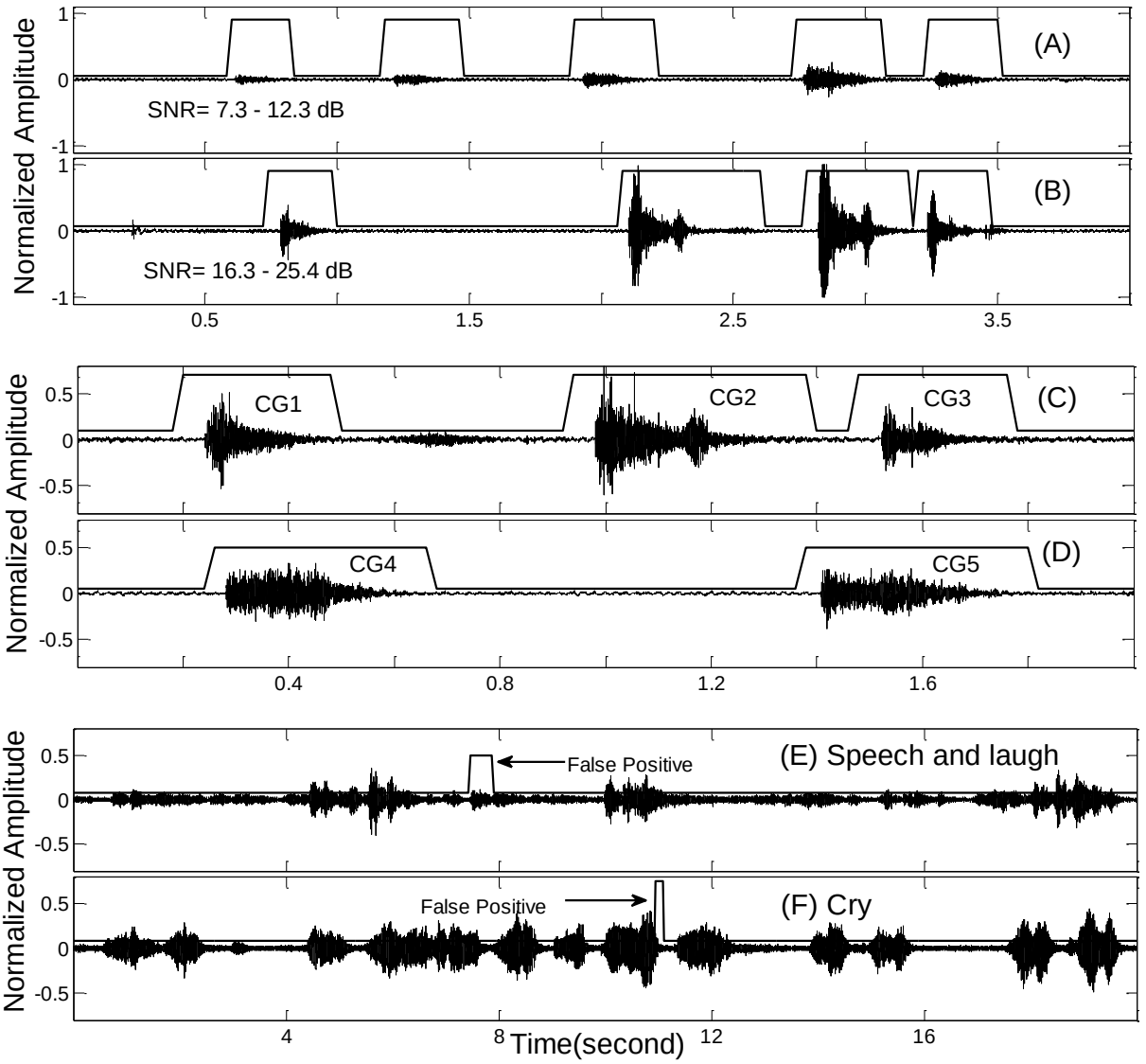

Figure 6.6: Illustration of cough segment identification using different types of coughs from the prospective study dataset PSD. It can be seen that the proposed algorithm is robust against: (i) SNR variation (graphs in (A) and (B)); (ii) waveform-shapes variations (graphs in (C) and (D)); and (iii) in rejecting non-cough sounds (graphs in (E) and F)).

In Figure 6.6, I further illustrate the performance of the proposed algorithm on sound recordings with different SNRs and a range of cough waveforms. All the sound segments in Figure 6.6 were taken from the prospective dataset PSD. In Figures 6.6(A) and 6.6(B), in spite of significant variations in cough intensity, the proposed algorithm is able to identify all the cough events. This signifies that the algorithm performance is robust against intensity variations. The waveform-shapes (e.g. the number of humps) of cough sounds can vary widely, especially in disease conditions. In order to test whether the algorithm is sensitive to this change, I tested it on two sound segments consisting of coughs with different wave-shapes (see Figures 6.6(C) and 6.6(D)). As can be seen from these figures, my algorithm identifies the entire cough segment irrespective of their wave-shape. Figures 6.6(E) and 6.6(F) show the performance of the algorithm on a typical sound segment consisting of speech, laughter, and cries. This result shows the robustness of the algorithm in discounting non-cough sounds, regardless of their intensity.

6.4. Discussion

In this study, I propose an automated method to identify cough events from paediatric sound recordings acquired with non-contact microphones in the natural environment of a respiratory ward. The subject cohort for my study was children from 3 to 71 months of age with respiratory diseases such as pneumonia, bronchitis, and nasopharyngitis. Working with the sound recordings from 24 paediatric subjects, I showed that my method was capable of identifying cough events with sensitivity and a specificity of > 90 percent. To the best of my knowledge, my work is the first ever attempt to identify cough events from the paediatric sound recordings.

All previous research on cough event identification is based on sound recordings from adult subjects. Yet, it is well known that the airway anatomy and functionality of children are quite different to those of adults and that cough sounds from the two groups have different characteristics.

The vast majority of my dataset consisted of coughs from pneumonia patients. This is a distinct novelty. The nature of the cough changes due to pneumonia; specifically, the coughs become smaller in magnitude, thus lowering the signal to noise ratio. Conventional detection methods that depend on features such as gradients and the number of humps in cough waveforms are not appropriate, due to the sharp decrease of the gradients and deteriorations in the hump structures. My method is novel in that it does not rely on intensity features alone, it does not depend on the number of humps, and it does not need calibration every time it is used. All of these features make my work unique.

Another novelty of my study is that I was able to analyse spontaneous coughs from children, whereas previous research methods [36, 113] only reported results on voluntary coughs in adults. Spontaneous coughs contain a wide range of natural variety in terms of cough sound intensity, duration, and waveform shape, which cannot be expected in a voluntary cough that could be consciously or subconsciously controlled by an adult subject. Furthermore, the spontaneous (and continuous) sound recordings capture other undesirable sound events such as toy noise, children's laughter and cries, and adult conversations. My algorithms assume total naiveté to the time of occurrence of any desirable sound event, whereas voluntary recording enables a reasonable control over undesirable sounds and the time of occurrence of desirable events. These differences make spontaneous coughs much more difficult to detect in a continuous sound stream.

I designed and tested my algorithm on data recorded using non-contact bedside microphones. Previous studies have used body-contact sensors to record sounds [34-36]. While contact sensors may provide an easier signal to work with, the need for a physical sensor to have contact with the patient is problematic. Specifically, contact sensors are inconvenient to use on paediatric subjects, and they cause a risk of cross-infection among pneumonia patients. These differences in measurement protocols make the results difficult to compare with those of previous studies.

6.5. Conclusion

The cough segmentation method proposed in this study achieved an accuracy, sensitivity, and specificity of 97.4 percent, 92.8 percent and 97.5 percent, respectively on a prospective dataset. The Kappa agreement between my method and the human observer was 0.65. Based on these results, I conclude that my method has the potential to be used in automated cough logger applications. It should also be possible to use it as the front-end of an automatic cough analysis system. The results presented in this study should be further verified with a larger dataset.

Novelty and impact:

- This is the first method facilitating automated cough segmentation for long-term recordings in paediatric populations.
- The method is non-contact, inexpensive, and reproducible.
- Can be used for a cough counter as well as the front end of a cough sound analysis system.

Chapter 7

Conclusion and Future Work

This thesis contributes to the development of novel methods for cough assessment and its application to screen respiratory diseases. It addresses three major gaps in the cough analysis fields: first, it develops an automated method for wet/dry cough identification; second, it develops a cough sound analysis based method that can be substituted for the bronchodilator test; and third, it develops an automated method for segmenting cough sounds from the continuous recordings.

Currently, wet/dry cough identification is manually identified by physicians during physical examinations. However, the perceptions of wet/dry cough differ among paediatricians due to their individual skills and experiences. In **Chapter 4,** I addressed the problem by proposing a novel method to identify wet/dry cough automatically. The proposed method achieved 84 percent sensitivity and 76 percent specificity, demonstrating that the method has potential to be implemented for cough quality monitoring especially in the home environment. It also facilitates the study of the wet/dry cough in long term recordings.

In **Chapter 5**, a cough sound analysis based method was proposed as an alternative to the bronchodilator test. Cough is the primary symptom of both pneumonia and asthma. Using non-contact sensors to obtain the cough signal facilitates the development of low cost diagnostic tools. Non-contact measurements reduce the risk of cross-infection and require less effort as no sterilization is required. All these benefits are useful to address the problems of pneumonia/asthma diagnosis in resource-limited settings. The proposed method achieved 100 percent sensitivity, 80 percent specificity, and 0.8 Kappa agreements in separating pneumonia from asthma subjects. The performances were computed from 20 subjects following the ‘leave one out’ validation method. The results show that analysis of cough sounds has the potential to be developed as a substitute for the bronchodilator test in resource-limited settings.

Chapter 6 proposed the development of an automated method for segmenting coughs from recordings obtained in paediatric populations. The proposed method achieved high performance with 93 percent sensitivity, 98 percent specificity and 0.65 percent Kappa. To the best of my knowledge, this is the first effort in this field. Previous research focused on the development of cough counting devices for the adult population and researchers relied on sound intensity features to

detect cough sounds. It is worth noting, however, that counting cough frequency is only one aspect of cough quantification study. There are numerous areas of cough quantification (i.e. wetness index, pneumonia cough index) that have not yet been explored. My method is intended to facilitate those purposes as it has the potential to be developed as an automated cough counting device, as well as the front end of cough analysis systems. Researchers and health workers no longer have to manually segment cough samples for the recordings.

The findings of this current research show that cough analysis can be carried automatically with minimum intervention from human operators. It also shows the potential for cough sound analysis as a means to develop a low cost, convenient, and accurate diagnostic tool for paediatric pneumonia/asthma. It can be augmented to suit the existing WHO guidelines for pneumonia management in the resource-limited settings.

It should be noted that the positive results from this thesis should be followed up with further studies before the proposed method is applied in the field.

7.1. Direction for future work

In this thesis, the cough samples were recorded using non-contact microphone sensors with a frequency range of 20 Hz to 20 kHz. It is understood that the propagation of sound signals correspond to their wavelength/frequency. To achieve the maximum performance, it is necessary to investigate the effect of distance variations between the sensors and the subjects and then to define the optimum distance to gain the maximum performance. Further, sounds recorded in the narrow room (e.g. paediatric wards in low-income countries) could produce echo that contributes to the recorded cough sounds. Studies on the effect of echo and an echo cancellation algorithm would be useful to improve the cough processing algorithm further.

The performance of the methods presented in this thesis can be improved by further fine tuning the parameters used in the algorithm. My methods are mostly inspired by the features and algorithms used in the speech processing field. Despite the positive results, an advanced signal-processing algorithm could be implemented to deepen the study on pneumonia/asthma and to obtain more definitive features/algorithms to separate those diseases.

Lung sounds such as crackles are commonly reported in pneumonia subjects. This information can be augmented to support the results from cough sound analysis. Further, crackles data can be used to validate the hypothesis that crackles present in the cough sounds of pneumonia subjects. If the hypothesis is true then non-contact detection of crackles via cough can be developed.

The database of cough sounds from paediatric subjects is extremely limited due to the lack of pertinent studies; therefore, developing a comprehensive database is one of the important tasks that should be completed. Studies on the influence of gender, age, and co-morbidities to the performance of the algorithm are only possible using a large database. Database augmentation is a key element in the development of robust algorithms. To this end, testing in an adequate number of prospective datasets is required before implementing the algorithm in the field.

One of the ultimate goals of this study is to develop pneumonia diagnosis tools for resource-limited settings. It is essential to develop the method into a portable device or mobile phone applications to enable production of low cost systems.

Bibliography

[1] C. Hsiao, D. K. Cherry, P. C. Beatty, and E. A. Rechsteiner, "National Ambulatory Medical Care Survey: 2007 Summary," *National Health Statistics Reports*, vol. 27, 2010.

[2] H. Britt, G. C. Muiller, J. Charles, J. Henderson, C. Bayram, Y. Pan, L. Valenti, C. Harrison, J. O'Halloran, and S. Fahridin, "General practice activity in Australia 2009-10," in *General Practice*, no. 27., cat no. GEP 27., Canberra: AIHW, 2010.

[3] World Health Organization, "WHO-recommended standards for surveillance of selected-preventable diseases," *World Health Organization*, Geneva, WHO/ V&B/03.01, 2003.

[4] C. L. French, R. S. Irwin, F. J. Curley, and C. J. Krikorian, "Impact of chronic cough on quality of life," *Arch. Intern. Med.*, vol. 158, pp. 1657 - 1661, 1998.

[5] W. Ma, L. Yu, Y. Wang, X. Li, H. Lu, and Z. Qiu, "Changes in health-related quality of life and clinical implications in Chinese patients with chronic cough," *Cough*, vol. 5, no. 7, 2009.

[6] L. McGarvey, C. Carton, L. Gamble, L. Heaney, R. Shepherd, M. Ennis, and J. MacMahon, "Prevalence of psychomorbidity among patients with chronic cough," *Cough*, vol. 2, no. 1, pp. 2-4, 2006.

[7] P. V. Dicpinigaitis, R. Tso, and G. Banauch, "Prevalence of depressive symptoms among patients with chronic cough," *Chest*, vol. 130, no. 6, pp. 1839-43, 2006.

[8] K. W. Purdy, J. W. Hay, M. F. Botteman, and J. I. Ward, "Evaluation of strategies for use of acellular pertussis vaccine in adolescents and adults: A cost-benefit analysis," *Clinical Infectious Diseases*, vol. 39, pp. 20-28, 2004.

[9] S. Hollinghurst, C. Gorst, T. Fahey, and A. Hay, "Measuring the financial burden of acute cough in pre-school children: A cost of illness study," *BMC Family Practice*, vol. 9, 2008.

[10] A. M. Fendrick, A. S. Monto, B. Nightengale, and M. Sarnes, "The economic burden of non-influenza-related viral respiratory tract infection in the United States," *Arch. Intern. Med.*, vol. 163, no. 4, pp. 487-94, 2003.

[11] R. S. Irwin, L. P. Boulet, M. M. Cloutier, R. Fuller, P. M. Gold, V. Hoffstein, A. J. Ing, F. D. McCool, P. O'Byrne, R. H. Poe, U. B. S. Prakash, M. R. Pratter, and B. K. Rubin, "Managing cough as a defense mechanism and as a symptom: A consensus panel report of the American college of chest physicians," *Chest*, vol. 114, no. 2 supplement, pp. 133S-181S, 1998.

[12] F. De Blasio, J. Virchow, M. Polverino, A. Zanasi, P. Behrakis, G. Kilinc, R. Balsamo, G. De Danieli, and L. Lanata, "Cough management: A practical approach," *Cough*, vol. 7, no 7, , 2011.

[13] A. J. Hotaling and G. T. Moynihan, "Cough," *Pediatric Otolaryngology*, vol. 2, Philadelphia: Saunders, 2003, pp. 1395-1404.

[14] P. G. Gibson, A. B. Chang, N. J. Glasgow, P. W. Holmes, P. Katelaris, A. S. Kemp, L. I. Landau, S. Mazzone, P. Newcombe, P. V. Asperen, and A. E. Vertigan, "CICADA: Cough in children and adults: diagnosis and assessment. Australian cough guidelines summary statement," *Med. J. Aust.*, vol. 192 (5), pp. 265-271, 2010.

[15] R. S. Irwin, M. H. Baumann, D. C. Bolser, L. P. Boulet, S. S. Braman, C. E. Brightling, K. K. Brown, B. J. Canning, A. B. Chang, P. V. Dicpinigaitis, R. Eccles, W. B. Glomb, L. B. Goldstein, L. M. Graham, F. E. Hargreave, P. A. Kvale, S. Z. Lewis, F. D. McCool, D. C. McCrory, U. B. Prakash, M. R. Pratter, M. J. Rosen, E. Schulman, J. J. Shannon, H. C. Smith, S. M. Tarlo, and American College of Chest Physicians (ACCP), "Diagnosis and

management of cough executive summary: ACCP evidence-based clinical practice guidelines," *Chest*, vol. 129, no. 1 supplement, pp. 1S-23S, 2006.
[16] J. Korpás, J. Sadlonová, and M. Vrabec, "Analysis of the cough sound: An overview," *Pulmonary Pharmacology*, vol. 9, pp. 261-268, 1996.
[17] J. A. Smith, H. Ashurst, S. Jack, A. Woodcock, and J. Earis, "The description of cough sounds by healthcare professionals," *Cough*, vol. 2, no. 1, 2006.
[18] Y. H. Hiew, J. A. Smith, J. E. Earis, B. M. G. Cheetham, and A. A. Woodcock, "DSP algorithm for cough identification and counting, *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* " vol. 4, pp. 3888-3891, 2002.
[19] J. Knocikova, J. Korpas, M. Vrabec, and M. Javorka, "Wavelet analysis of voluntary cough sound in patients with respiratory diseases," *Journal of Physiology and Pharmacology*: An official journal of the Polish Physiological Society, vol. 59, pp. 331-40., 2009.
[20] K. F. Chung, "Measurement of cough," *Respiratory Physiology & Neurobiology*, vol. 152, pp. 329-339, 2006.
[21] G. A. Fontana and J. Widdicombe, "What is cough and what should be measured?," *Pulmonary Pharmacology & Therapeutics*, vol. 20, pp. 307-312, 2007.
[22] J. A. Smith, "Assessing efficacy of therapy for cough," *Otolaryngol Clin. North Am.*, vol. 43, no. 1, pp. 157-66, 2010.
[23] J. A. Smith and A. Woodcock, "New developments in the objective assessment of cough," *Lung*, vol. 186, pp. 48-54, 2008.
[24] A. M. Li, C. Lex, A. Zacharasiewicz, E. Wong, E. Erin, T. Hansel, N. M. Wilson, and A. Bush, "Cough frequency in children with stable asthma: Correlation with lung function, exhaled nitric oxide, and sputum eosinophil count," *Thorax*, vol. 58, no. 11, pp. 974-978, 2003.
[25] A. M. Li, T. W. T. Tsang, D. F. Y. Chan, H. S. Lam, H. K. So, R. Y. T. Sung, and T. F. Fok, "Cough frequency in children with mild asthma correlates with sputum neutrophil count," *Thorax*, vol. 61, pp. 747-750, September 2006.
[26] A. H. Thomson, C. Pratt, and H. Simpson, "Nocturnal cough in asthma," *Archives of Disease in Childhood*, vol. 62, no. 10, pp. 1001–1004, 1987.
[27] D. Ginsberg, "An unidentified monster in the bed: Assessing nocturnal asthma in children," *Mcgill J. Med.*, vol. 12, no. 1, pp. 31–38, 2009.
[28] C. F. Everett and A. H. Morice, "Clinical history in gastroesophageal cough," *Respiratory Medicine*, vol. 101, pp. 345-348, 2007.
[29] M. Jinnai, A. Niimi, M. Takemura, H. Matsumoto, Y. Konda, and M. Mishima, "Gastroesophageal reflux-associated chronic cough in an adolescent and the diagnostic implications: A case report," *Cough*, vol. 4, no. 5, 2008.
[30] P. V. Dicpinigaitis, G. Colice, M. Goolsby, G. Rogg, S. Spector, and B. Winther, "Acute cough: A diagnostic and therapeutic challenge," *Cough*, vol. 5, no. 11, 2009.
[31] J. M. Marchant, I. B. Masters, S. M. Taylor, and A. B. Chang, "Utility of signs and symptoms of chronic cough in predicting specific cause in children," *Thorax*, vol. 61, pp. 694-698, 2006.
[32] S. Barry, A. Dane, A. Morice, and A. Walmsley, "The automatic recognition and counting of cough," *Cough*, vol. 2, no. 1, p. 8, 2006.
[33] S. Matos, S. S. Birring, I. D. Pavord, and D. H. Evans, "An automated system for 24-h monitoring of cough frequency: The Leicester cough monitor," *IEEE Trans. on Biomed. Eng.*, vol. 54, pp. 1472-1479, 2007.
[34] M. A. Coyle, D. B. Keenan, L. S. Henderson, M. L. Watkins, B. K. Haumann, D. W. Mayleben, and M. G. Wilson, "Evaluation of an ambulatory system for the quantification of cough frequency in patients with chronic obstructive pulmonary disease," *Cough*, vol. 1, p. 3, 2005.

[35] K. McGuinness, J. Morris, J. Lowe, A. Woodcock, and J. A. Smith, "Automated cough detection: A novel approach [abstract]," *Am. J. Resp. Crit. Care Med.*, p. 175: A381, 2007.
[36] E. Vizel, M. Yigla, Y. Goryachev, E. Dekel, V. Felis, H. Levi, I. Kroin, S. Godfrey, and N. Gavriely, "Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions," *Cough*, vol. 6, no. 1, p. 3, 2010.
[37] A. B. Chang, "Pediatric cough: Children are not miniature adults," *Lung*, vol. 188, no. 1, pp. 33-40, 2010.
[38] S. S. Birring, B. Prudon, A. J. Carr, S. J. Singh, M. D. L. Morgan, and I. D. Pavord, "Development of a symptom specific health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ)," *Thorax*, vol. 58, pp. 339-343, 2003.
[39] C. T. French, R. S. Irwin, K. E. Fletcher, and T. M. Adams, "Evaluation of a cough-specific quality-of-life questionnaire," *Chest*, vol. 121, no. 4, pp. 1123-31, 2002.
[40] P. A. Newcombe, J. K. Sheffield, E. F. Juniper, H. L. Petsky, C. Willis, and A. B. Chang, "Validation of a parent-proxy quality of life questionnaire for paediatric chronic cough (PC-QOL)," *Thorax*, vol. 65(9), pp. 819-23, 2010.
[41] I. Baiardini, F. Braido, O. Fassio, F. Tarantini, M. Pasquali, F. Tarchino, A. Berlendis, and G. W. Canonica, "A new tool to assess and monitor the burden of chronic cough on quality of life: Chronic Cough Impact Questionnaire," *Allergy*, vol. 60, no. 4, pp. 482-8, 2005.
[42] A. Gulsvik and O. K. Refvem, "A scoring system on respiratory symptoms," *European Respiratory Journal*, vol. 1, no. 5, pp. 428-432, 1988.
[43] C. E. Brightling, W. Monterio, R. H. Green, D. Parker, M. D. Morgan, A. J. Wardlaw, and D. Pavord, "Induced sputum and other outcome measures in chronic obstructive pulmonary disease: safety and repeatability," *Respir. Med.*, vol. 95, no. 12, pp. 999-1002, 2001.
[44] A. Falconer, C. Oldman, and P. Helms, "Poor agreement between reported and recorded nocturnal cough in asthma," *Pediatric Pulmonology*, vol. 15, no. 4, pp. 209-11, 1993.
[45] A. M. Brooke, P. C. Lambert, P. R. Burton, C. Clarke, D. K. Luyt, and H. Simpson, "Night cough in a population-based sample of children: Characteristics, relation to symptoms and associations with measures of asthma severity," *Eur. Respir. J.*, vol. 9, no. 1, pp. 65-71, 1996.
[46] R. E. Dales, J. White, C. Bhumgara, and E. McMullen, "Parental reporting of childrens' coughing is biased," *European Journal of Epidemiology*, vol. 13, pp. 541-545, 1997.
[47] A. B. Chang, L. I. Landau, P. P. Van Asperen, N. J. Glasgow, C. F. Robertson, J. M. Marchant, and C. M. Mellis, "Cough in children: definitions and clinical evaluation," *Med J. Aust.*, vol. 184, pp. 398-403, 2006.
[48] B. A. Cunha, *Pneumonia Essentials*, 3rd ed., Sudbury, Mass.: Physicians' Press, 2010.
[49] A. B. Chang, M. Eastburn, J. Faoagali, and I. Masters, "Cough quality in children: A comparison of subjective vs. bronchoscopic findings," *Respir. Res.*, vol. 6, 2005.
[50] A. B. Chang, "Causes of cough, assessment and measurement in children," in *Cough: Mechanisms, causes and therapy*, Oxford: Blackwell Science, 2003.
[51] A. B. Chang, P. D. Phelan, S. M. Sawyer, S. Del Brocco, and C. F. Robertson, "Cough sensitivity in children with asthma, recurrent cough, and cystic fibrosis," *Archives of Disease in Childhood*, vol. 77, pp. 331-4, 1997.
[52] A. B. Chang, P. D. Phelan, S. M. Sawyer, and C. F. Robertson, "Airway hyper-responsiveness and cough-receptor sensitivity in children with recurrent cough," *Am. J. Respir. Crit. Care Med.*, vol. 155, pp. 1935-9, 1997.
[53] C. L. F. Walker, I. Rudan, L. Liu, H. Nair, E. Theodoratou, Z. A. Bhutta, K. L. O'Brien, H. Campbell, and R. E. Black, "Global burden of childhood pneumonia and diarrhoea," *The Lancet*, vol. 381, pp. 1405-1416, 2013.
[54] R. E. Black, S. S. Morris, and J. Bryce, "Where and why are 10 million children dying every year?" *The Lancet*, vol. 361, pp. 2226-2234, 2003.

[55] World Health Organization, "Handbook: IMCI integrated management of childhood illness," Geneva: World Health Organization and UNICEF, 2005.
[56] C. Horwood, K. Vermaak, N. Rollins, L. Haskins, P. Nkosi, and S. Qazi, "An evaluation of the quality of IMCI assessments among IMCI trained health workers in South Africa," *PLoS ONE*, vol. 4, p. e5937, 2009.
[57] S. J. Armstrong, J. Bryce, D. de Savigny, T. Lambrechts, C. Mbuya, L. Mgalula, K. Wilczynska, and Tanzania IMCI Multi-Country Evaluation Health Facility Survey Study Group, "The effect of Integrated Management of Childhood Illness on observed quality of care of under-fives in rural Tanzania," *Health Policy and Planning*, vol. 19, pp. 1-10, 2004.
[58] D. Shah and H. P. Sachdev, "Evaluation of the WHO/UNICEF algorithm for integrated management of childhood illness between the age of two months to five years," *Indian Pediatr.*, vol. 36, no. 8, p. 767, 1999.
[59] R. Nantanda, J. K. Tumwine, G. Ndeezi, and M. S. Ostergaard, "Asthma and pneumonia among children less than five years with acute respiratory symptoms in Mulago Hospital, Uganda: Evidence of under-diagnosis of asthma," *PLoS ONE*, vol. 8, p. e81562, 2013.
[60] H. P. Sachdec, B. Vasanthi, L. Satyanarayana, and R. K. Puri, "Simple predictors to differentiate acute asthma from ARI in children: Implications for refining case management in the ARI Control Programme," *Indian Pediatr.*, vol. 31, no. 10, pp. 1251-9, 1994.
[61] H. P. Sachdec, S. C. Mahajan, and A. Garg, "Improving antibiotic and bronchodilator prescription in children presenting with difficult breathing: Experience from urban hospital in India," *Indian Pediatr.*, vol. 38, no. 8, pp. 827-38, 2001.
[62] Z. A. Bhutta, "Childhood pneumonia in developing countries," *BMJ*, vol. 333, no. 7569, pp. 612-13, 2006.
[63] M. A. Cardoso, C. M. Nascimento-Carvalho, F. Ferrero, F. M. Alves, and S. N. Cousens, "Adding fever to WHO criteria for diagnosing pneumonia enhances the ability to identify pneumonia cases among wheezing children," *Archives of Disease in Childhood*, vol. 96, no. 1, pp. 58-61, 2011.
[64] A. Falade, H. Tschappeler, B. M. Greenwood, and K. Mulholland, "Use of simple clinical signs to predict pneumonia in young Gambian children: the influence of malnutrition," *Bulletin of the World Health Organization*, vol. 73, no. 3, p. 299, 1995.
[65] M. Harari, V. Spooner, S. Meisner, M. Carney, F. Shann, and J. de Campo, "Clinical signs of pneumonia in children," *The Lancet*, vol. 338, pp. 928-930, 1991.
[66] M. Palafox, H. Guiscafré, H. Reyes, O. Munoz, and H. Martínez, "Diagnostic value of tachypnoea in pneumonia defined radiologically," *Archives of Disease in Childhood*, vol. 82, no. 1, pp. 41-5, 2000.
[67] World Health Organization, "Technical updates of guidelines on Integrated Management of Childhood Illness (IMCI): Evidence and recommendations for further adaptions," Geneva: World Health Organization, 2005.
[68] T. Hazir, S. Qazi, Y. B. Nisar, S. Ansari, S. Maqbool, S. Randhawa, Z. Kundi, R. Asghar, and S. Aslam, "Assessment and management of children aged 1-59 months presenting with wheeze, fast breathing, and/or lower chest indrawing: Results of a multicentre descriptive study in Pakistan," *Archives of Disease in Childhood*, vol. 89, pp. 1049-1054, 2004.
[69] K. H. Naifeh, "Basic anatomy and physiology of the respiratory system and the autonomic nervous system," in *Behavioral and Physchological Approaches to Breathing Disorder*, Timmons B.H. and Ley R., eds., New York: Plennum Press, 1994.
[70] N. C. Staub, *Basic Respiratory Physiology*. New York: Curchill Livingstone, 1991.
[71] S. Sircar, *Principles of Medical Physiology*. New York: Thieme, 2008.
[72] E. N. Marieb, *Essentials of Human Anatomy and Physiology*, 9th ed., San Fransisco: Pearson, 2009.
[73] J. B. West, *Respiratory Physiology: The Essentials*. Philadelphia: Lippincott Williams & Wilkins, 2005.

[74] D. D. Chiras, *Human Biology*, 8th ed., Burlington: Jones & Barlett Learning, 2014.
[75] K. Rogers, *The Respiratory System*, USA: Britannica Educational Publishing, 2011.
[76] Springhouse, Lippincott Professional *Guides: Anatomy & Physiology*, 2nd ed., USA: Lippincott William & Wilkins, 2002.
[77] L. I. Landau and L. M. Taussig, "Early childhood origins and economic impact of respiratory disease throughout life," in *Pediatric Respiratory Medicine*, L. M. Taussig, L. I. Landau, P. N. Le Souef, W. J. Morgan, F. D. Martinez, and P. D. Sly, Eds., 2nd ed., Philadelphia: Mosby Elseiver, 2008.
[78] S. J. Bourke, and G. P. Burns, *Lecture Notes: Respiratory Medicine*, vol. 35, John Wiley & Sons, 2011.
[79] H. Patel, *Respiratory System*, 3rd ed., Edinburgh: Mosby Elsevier, 2007.
[80] P. D. Phelan, A. Olinsky, and C. F. Robertson, *Respiratory Illness in Children*, 4th ed., Oxford: Blackwell Scientific Publications, 1994.
[81] C. M. Alper, E. M. Arjmand, S. E. Stool, and M. L. Casselbrant, *Pediatric otolaryngology*, 4th ed., vol. 1, Saunders, 2002.
[82] V. Krishna, *Textbook of Pathology*, Hyderabad: Orient Longman Private Ltd, 2004.
[83] F. Blasi, S. Aliberti, M. Pappalettera, and P. Tarsia, "100 years of respiratory medicine: Pneumonia," *Respir. Med.*, vol. 101, no. 5, pp. 875-881, 2007.
[84] G. G. Burton and T. R. Jarins, *Clinical Manifestations and Assessment of Respiratory Disease*, 6th ed. Maryland Heights, MO: Mosby/Elseiver, 2011.
[85] I. Rudan, C. Boschi-Pinto, Z. Biloglav, K. Mulholland, and H. Campbell, "Epidemiology and etiology of childhood pneumonia," *Bulletin of the World Health Organization*, vol. 86, pp. 408-416, 2008.
[86] European Respiratory Society, *European Lung White: Pneumonia*, 2nd ed., Sheffield, UK: European Respiratory Society/European Lung Foundation, 2003.
[87] G. Scott, H. Scott, M. Turley, and M. Baker, "Economic cost of community-acquired pneumonia in New Zealand adults," *Journal of New Zealand Medical Association*, vol. 117, 2004.
[88] Australian Bureau of Statistics, "Census of population and housing," *Australian Bureau of Statistics*, Canberra, 2001.
[89] A. Bush and J. C. Davies, *An Atlas of Investigation and Management: Paediatric Respiratory Diseases, Airways, and Infection*, Oxford: Clinical Publishing, 2011.
[90] R. Lichenstein, A. H. Suggs, and J. Campbell, "Pediatric pneumonia," *Emerg. Med. Clin. N. Am.*, vol. 21, no. 2, pp. 437-451, 2003.
[91] S. Block, J. Hedrick, M. R. Hammerschlag, G. H. Casell, and J. C. Craft, "Mycoplasma pneumoniae and chlamydia pneumoniae in pediatric community-acquired pneumonia: Comparative efficacy and safety of clarithromycin vs. erythromycin ethylsuccinate," *Pediatr. Infect. Dis. J.*, vol. 14, no. 6, pp. 471-477, 1995.
[92] T. Heiskanen-Kosma, M. Korppi, and C. Jokinen, "Etiology of childhood pneumonia: Serologic results of a prospective, population-based study," *Pediatr. Infect. Dis. J.*, vol. 17, pp. 986-91, 1998.
[93] L. Wubbel, L. Muniz, and A. Ahmed, "Etiology and treatment of community-acquired pneumonia in ambulatory children," *Pediatr. Infect. Dis. J.*, vol. 18, pp. 98-104, 1999.
[94] C. G. Victora, B. R. Kirkwood, A. Ashworth, R. E. Black, S. Rogers, S. Sawazal, H. Campbell, and S. Gove, "Potential of interventions for the prevention of childhood pneumonia in developing countries: improving nutrition," *Am. J. Clin. Nutr.*, vol. 70, pp. 309-20, 1999.
[95] R. B. Yellanthoor and V. K. B. Shah, "Prevalence of malnutrition among under-five year old children with acute lower respiratory tract infection hospitalized at Udupi District Hospital," *Arch. Pediatr. Infect. Dis.*, vol. 5, pp. 203-206, 2013.

[96] S. I. Pelton and M. R. Hammerschlag, "Overcoming current obstacles in the management of bacterial community-acquired pneumonia in ambulatory children," *Clin. Pediatr. (Phila.)*, vol. 44, no. 1, pp. 1-17, 2005.
[97] World Health Organization, "WHO Fact sheet N°331: Pneumonia," World Health Organization, 2013.
[98] S. Farzan, *A Concise Handbook of Respiratory Diseases*. Connecticut, US: Apple & Lange, 1997.
[99] T. J. Marrie, C. Y. Lau, S. L. Wheeler, C. J. Wong, and B. G. Feagan, "Predictors of symptom resolution in patients with community-acquired pneumonia," *Clinical Infectious Diseases*, vol. 31, pp. 1362-1367, 2000.
[100] R. L. H. Murphy, A. Vyshedskiy, V. A. Power-Charnitsky, Verna-Ann, D. S. Bana, P. M. Marinelli, A. Wong-Tse, and R. Paciej, "Automated lung sound analysis in patients with pneumonia," *Respiratory Care*, vol. 49, pp. 1490-1497, 2004.
[101] P. Piirila and A. Sovijarvi, "Crackles: Recording, analysis and clinical significance," *European Respiratory Journal*, vol. 8, pp. 2139-2148, 1995.
[102] I. C. Michelow, K. Olsen, J. Lozano, N. K. Rollins, L. B. Duffy, T. Ziegler, J. Kauppila, M. Leinonen, and G. H. McCracken, "Epidemiology and clinical characteristics of community-acquired pneumonia in hospitalized children," *Pediatrics*, vol. 113, pp. 701-707, 2004.
[103] D. M. Hansell, D. A. Lynch, H. P. McAdams, and A. A. Bankier, *Imaging of Diseases of the Chest*, 5th ed., Edinburgh: Mosby, 2010.
[104] R. Izadnegahdar, A. L. Cohen, K. P. Klugman, and S. A. Qazi, "Childhood pneumonia in developing countries," *The Lancet Respiratory Medicine*, vol. 1, pp. 574-584, 2013.
[105] S. M. Graham, M. English, T. Hazir, P. Enarson, and T. Duke, "Challenges to improving case management of childhood pneumonia at health facilities in resource-limited settings," *Bulletin of the World Health Organization*, vol. 86, no. 5, pp. 349-55, 2008.
[106] S. Farzan, "Cough and sputum production," in *Clinical Methods: The History, Physical, and Laboratory Examinations*, H. K. Walker, W. D. Hall, and J. W. Hurst, Eds., 3rd ed., Atlanta: Butterworths, 1990.
[107] A. H. Morice, G. A. Fontana, M. G. Belvisi, S. S. Birring, K. F. Chung, P. V. Dicpinigaitis, J. A. Kastelik, L. P. McGarvey, J. A. Smith, M. Tatar, and J. Widdicombe, "ERS guidelines on the assessment of cough," *Eur. Respir. J.*, vol. 29, pp. 1256-76, 2007.
[108] J. Korpás and Z. Tomori, *Cough and Other Respiratory Reflexes*, Karger, 1979.
[109] J. Mallol, J. Crane, E. von Mutius, J. Odhiambo, U. Keil, and A. Stewart, "The International Study of Asthma and Allergies in Childhood (ISAAC) phase three: A global synthesis," *Allergologia et Immunopathologia*, vol. 41, pp. 73-85, 2013.
[110] A. B. Goldsobel and B. E. Chipps, "Cough in the pediatric population," *The Journal of Pediatrics*, vol. 156, pp. 352-358, 2010.
[111] A. B. Chang and W. B. Glomb, "Guidelines for evaluating chronic cough in pediatrics : ACCP evidence-based clinical practice guidelines," *Chest*, vol. 129, no. 1 supplement, pp. 260S-283S, 2006.
[112] J. Martinek, M. Tatar, and M. Javorka, "Distinction between voluntary cough sound and speech in volunteers by spectral and complexity analysis," *J. of Physiol. and Pharma.*, vol. 59 p. 433, 2008.
[113] S. H. Shin, T. Hashimoto, and S. Hatano, "Automatic detection system for cough sounds as a symptom of abnormal health condition," *IEEE Trans. on Inf. Technol. Biomed.*, vol. 13, pp. 486-493, 2009.
[114] P. Piirila and A. R. Sovijarvi, "Differences in acoustic and dynamic characteristics of spontaneous cough in pulmonary diseases," *Chest*, vol. 96, pp. 46-53, 1989.
[115] U. R. Abeyratne, V. Swarnkar, R. Triasih, and A. Setyati, "Cough sound analysis can rapidly diagnose childhood pneumonia," *Annals of Biomedical Engineering*, vol. 41, no. 11, pp. 2448-62, 2013.

[116] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Trans. Speech Audio Process.*, vol. 9, pp. 504-512, 2001.
[117] M. E. Soto-Quiros, M. Soto-Martinez, and L. A. Hanson, "Epidemiological studies of the very high prevalence of asthma and related symptoms among school children in Costa Rica from 1989 to 1998," *Pediatr. Allergy Immunol.*, vol. 13, no. 5, pp. 342-9, 2002.
[118] J. D. Spengler, J. J. Jaakkola, H. Parise, B. A. Katsnelson, L. I. Privalova, and A. A. Kosheleva, "Housing characteristics and children's respiratory health in the Russian Federation," *Am. J. Public Health*, vol. 94, no. 4, pp. 657-62, 2004.
[119] D. Zgherea, S. Pagala, M. Mendiratta, M. G. Marcus, S. P. Shelov, and M. Kazachkov, "Bronchoscopic findings in children with chronic wet cough," *Pediatrics*, vol. 129, no. 2, pp. e364-9, 2012.
[120] S. Ishikawa, R. Kasparian, J. Blanco, D. Sotherland, R. Clubb, L. Kenny, J. Workowics, and K. MacDonell, "Observer variability in interpretation of voluntary cough of bronchitis and asthmatics," in *ILSA Proceedings*, Paris, 1987.
[121] A. Murata, "Discrimination of productive and non-productive cough by sound analysis," *Internal Medicine*, vol. 37, no. 9, pp. 732-735, 1998.
[122] H. Chatrzarrin, A. Arcelus, R. Goubran, and F. Knoefel, "Feature extraction for the differentiation of dry and wet cough sounds," in *IEEE International Workshop on Medical Measurements and Applications Proceedings (MeMeA)*, 2011, pp. 162-166.
[123] K. Rosenberry, W. T. Goldsmith, J. S. Reynolds, W. McKinney, and D. G. Frazer, "Gender differences in voluntary cough sound spectra demonstrated by an inverse power law analysis," in *24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society, Houston*, 2002, pp. 222-223.
[124] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A K-Means Clustering Algorithm," *Journal of Royal Statistical Society*, vol. 28, no. 1, pp. 100-108, 1979.
[125] C. W. Thorpe, L. J. Toop, and K. P. Dawson, "Towards a quantitative description of asthmatic cough sounds," *Eur. Respir. J.* , vol. 5, no. 6, pp. 685-92, 1992.
[126] W. Thorpe, M. Kurver, G. King, and S. Salome, "Acoustic analysis of cough sound," presented at *the Seven Australia and New Zealand Intelligent Information Systems Conf.*, Perth, Western Australia, 2001.
[127] A. B. Chang, G. J. Redding, and M. L. Everard, "Chronic wet cough: Protracted bronchitis, chronic suppurative lung disease and bronchiectasis," *Pediatr. Pulmonol.*, vol. 43, no. 6, pp. 519-31, 2008.
[128] World Health Organization, "Chronic respiratory diseases," Geneva: World Health Organization, 2012.
[129] J. Mallol, "Childhood asthma in developing countries. Low income aspects and related matters," *Allergologia et Immunophatologia*, vol. 28, no. 5, pp. 283-286, 2000.
[130] World Health Organization, "IMCI: Integrated Management of Childhood Illnesses," in Management of the child with a serious infection or severe malnutrition: Guidelines for care at the first-referral level in developing countries, ed: World Health Organization and UNICEF, 2000.
[131] D. P. Johns and R. Pierce, "Spirometry: the measurement and interpretation of ventilatory function in clinical practise," *National Asthma Council Australia*, Melbourne, 2008.
[132] L. J. Toop, C. W. Thorpe, and R. Fright, "Cough sound analysis: a new tool for the diagnosis of asthma?," *Fam Pract*, vol. 6(2), pp. 83-5, 1989.
[133] X. Huang, A. Acero, and H. Hon, *Spoken language processing: a guide to theory, algorithm and system development*. Upper Saddle River: Prentice Hall, 2001.
[134] L. E. Baum and J. A. Eagon, "An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to model for ecology," *Bull. Amer. Math. Soc.*, vol. 73(3), pp. 360-63, 1967.

[135] A. A. Woodcock, J. A. Smith, and K. McGuinnes, "Method for generating output data," United States Patent 7761302, 2010.
[136] S. Matos, S. S. Birring, I. D. Pavord, and H. Evans, "Detection of cough signals in continuous audio recordings using hidden Markov models," *IEEE Trans. on Biomed. Eng.*, vol. 53(6), pp. 1078-1083, 2006.
[137] M. A. Coyle, D. B. Keenan, A. Derchak, M. Sackner, F. Wilhelm, K. Gilroy, M. Gummels, D. M. Inma, P. Kennedy, M. Mitchnick, and A. Behar, "Systems and methods for respiratory event detection," United States Patent 7267652, 2007.
[138] N. Graviely, "Phonopneumograph system," United States Patent 6168568, 2001.
[139] N. Gavriely, M. Nissan, D. W. Cugell, and A. H. Rubin, "Respiratory health screening using pulmonary function tests and lung sound analysis," *Eur. Respir. J.*, vol. 7, pp. 35-42, January 1, 1994 1994.
[140] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang, "Phoneme recognition using time-delay neural networks," *IEEE Trans. Acoust.*, Speech Signal Process., vol. 37, pp. 328-339, 1989.
[141] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *IJCNN International Joint Conference on Neural Networks* 1990, vol. 3, pp. 21-26.
[142] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," in *IEEE International Conference on Neural Networks* 1993, pp. 586-591, vol. 1.
[143] S. Greenberg, W. A. Ainsworth, and R. R. Fay, *Speech Processing in Auditory System*. New York: Springer, 2004.
[144] E. Goldshtein, A. Tarasiuk, and Y. Zigel, "Automatic detection of obstructive sleep apnea using speech signals," *IEEE Trans. Biomed Eng. OnLine*, vol. 58, pp. 1373-1382, 2011.
[145] A. S. Karunajeewa, U. R. Abeyratne, and C. Hukins, "Multi-feature snore sound analysis in obstructive sleep apnea–hypopnea syndrome," *Physiological Measurement*, vol. 32, p. 83, 2011.
[146] H. Ghaemmaghami, U. R. Abeyratne, and C. Hukins, "Normal probability testing of snore signals for diagnosis of obstructive sleep apnea," in *Engineering in Medicine and Biology Society*, 2009, EMBC 2009, Annual International Conference of the IEEE, 2009, pp. 5551-5554.
[147] T. Lee, U. R. Abeyratne, K. Puvanendran, and K. Goh, "Formant structure and phase-coupling analysis of human snoring sounds for detection of obstructive sleep apnea," presented at the *Computer Methods in Biomechanics and Biomedical Engineering Conference*, 2000.
[148] A. K. Ng, T. S. Koh, E. Baey, T. H. Lee, U. R. Abeyratne, and K. Purvanendran, "Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea?," *Sleep Medicine*, vol. 9, pp. 894-98, 2008.
[149] J. E. Markel and A. H. Gray, *Linear prediction of speech*. New York: Springer-Verlag Inc, 1982.
[150] M. M. Sondhi, "New methods of pitch extraction," *IEEE Trans. Audio Electroacoust.*, vol. 16, pp. 262-266, 1968.
[151] U. R. Abeyratne, "Blind reconstruction of non-minimum-phase systems from 1-D oblique slices of bispectrum," *IEEE Proceedings Vision, Image and Signal Processing*, vol. 146, pp. 253-264, 1999.
[152] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, pp. 278-305, 1991.

Appendix A

A.1. Computed features

Let $\hat{s}_k[n]$ represents the k^{th} (k = 1, 2, …, K) sub-block/sub-segment having length N from signal $\hat{s}[n]$. The following features can be computed in each $\hat{s}_k[n]$.

A1.1. Mel-frequency cepstral coefficients (MFCC)

MFCC is widely used in speech processing [143]. It was found to be highly useful for snore analysis [144, 145] as well as cough analysis [33]. To compute the MFCCs, $\hat{s}_k[n]$ is processed through four successive stages: signal windowing, discrete Fourier transform, Mel-frequency filter banks filtering, and discrete cosine transform [133]. The equations for the Hamming window and discrete Fourier transform are given in (A-1) and (A-2) respectively.

$$w[n] = 0.54 - 0.46\cos(n\pi / N) \tag{A-1}$$

$$\zeta_k(\wp) = \sum_{n=0}^{N-1} \hat{s}_k[n] w[n] e^{-\frac{j\pi\wp n}{N}}, \qquad 0 \le \wp < N \tag{A-2}$$

Next, the magnitude of $\zeta_k(\wp)$ is passed to Mel filter banks where the output given in (A-3).

$$S[m] = \ln\left[\sum_{\wp=0}^{N-1} \left|\zeta_k(\wp)\right|^2 H_m[\wp]\right], \qquad 0 \le m \le M \tag{A-3}$$

where m is the number of Mel filter banks, which can be varied from 24 to 40. The transfer function of the Mel-filter is given in (A-4).

$$H_m[\wp] \begin{cases} 0 & , \wp < f(m\text{-}1) \\ \dfrac{(\wp - f[m-1])}{(f[m] - f[m-1])} & , \text{for } f[m\text{-}1[\le \wp \le f[m] \\ \dfrac{(f[m+1] - \wp)}{(f[m+1] - f[m])} & , \text{for } f[m] \le \wp \le l_{fc}(m+1) \\ 0 & , \text{for } \wp \ge f[m+1] \end{cases} \tag{A-4}$$

Let f_l and f_h be the lowest and the highest frequencies of the filter bank in Hz, and f_s are the sampling rate in Hz. The boundary points $f[m]$ are given as follows.

$$f[m] = \left(\frac{N}{f_s}\right) B^{-1}\left(B(f_l) + \frac{B(f_h) - B(f_l)}{M+1}\right) \tag{A-5}$$

The Mel-scale B and its inverse (B^{-1}) given in (A-6) and (A-7) respectively.

$$B(f) = 1125\ln(1 + f/700) \tag{A-6}$$

$$B^{-1}(m) = 700\left(\exp\left(m/1125\right) - 1\right) \tag{A-7}$$

Finally, the MFCCs are computed is computed using a discrete cosine transform as in (A-8).

$$\Phi[\eta] = \sum_{m=0}^{M-1} S[m]\cos\left(\frac{\pi\eta(m + 1/2)}{M}\right) \tag{A-8}$$

where η represents the number of cesptral coefficients. The first η =13 cesptral coefficients were used in the automatic segmentation work. In pneumonia and asthma classification using HMM, 13 of the first order differential ($\Delta\Phi$) and 13 of the second order differential ($\Delta\Delta\Phi$) of MFCC were augmented to capture the temporal signal dynamic.

A1.2. Non-Gaussianity score (NGS)

The NGS provides an easy method to quantify the deviation of a given signal from a Gaussian model. In a study on snore sound analysis [146], this feature showed a capability to screen obstructive sleep apnoea. In this work, the NGS is computed as follows. Let p be a model Gaussian probability plot denoted as (A-9). I compute γ, the inverse (F^{-1}) of normal Cumulative Distribution Function (*cdf*) at given probability ϑ denoted in (A-10).

$$\vartheta_i \begin{cases} 1 - 0.5^{1/N} & , \text{for } i = 1 \\ 0.5^{1/N} & , \text{for } i = N \\ i - 0.3175/N + 0.365 & , \text{otherwise} \end{cases} \tag{A-9}$$

$$\gamma = F^{-1}(\vartheta \mid \mu, \sigma) = \{\gamma : F(\gamma \mid \mu, \sigma) = \vartheta\} \tag{A-10}$$

where the mean (μ = *0)* and standard deviation (σ = 1). Suppose *r is* the discrete signal in $\hat{s}_k[n]$. The estimation of inverse normal *cdf* of $s_k[n]$ is given in (A-11) and the NGS in (A-12).

$$\tilde{\gamma} = \frac{\chi_3^{\gamma} - \chi_1^{\gamma}}{\chi_3^{x} - \chi_1^{x}}(r - \mu_x) \tag{A-11}$$

$$\psi_k = 1 - \frac{\sum_{n=1}^{N}(\tilde{\gamma} - \gamma)^2}{\sum_{n=1}^{N}(\tilde{\gamma} - \bar{\tilde{\gamma}})^2} \tag{A-12}$$

where χ_1 and χ_3 respectively is the first and third quartile of r.

A1.3. Zero crossing rate

The ZCR, defined as the total times a signal crosses the zero axis, is a simple but useful method to detect the periodic nature of a signal regardless of its magnitude. The ZCR feature Z_k is computed as follows.

$$Z_k = \frac{1}{N-1}\sum_{n=1}^{N-1}\Pi\left[\hat{s}_k(n)\hat{s}_k(n-1) < 0\right] \quad \text{(A-13)}$$

where the indicator function $\Pi[A]$ is *1* if the argument A is true and *0* for otherwise.

A1.4. Shannon entropy

Cough sound is a complex signal, which represents contributions from various sub-structures of the respiratory tract. Some of these components display pseudo-periodic structures, while others have a random stochastic character. In this work, I computed the Shannon entropy to capture these features. The Shannon entropy ($\hbar_k$) of a sub-block $\hat{s}_k[n]$ was obtained using definition in (A-14).

$$\hbar_k = -\sum_{n=1}^{N-1}\left(\hat{s}_k(n)^2\right)\ln\left(\hat{s}_k(n)^2\right), \quad 1 \le n \le N\text{-}1 \quad \text{(A-14)}$$

A1.5. Formant frequency

In speech, Formant frequencies show the characteristics of vocal tract resonances; in snore sound analysis they indicate the resonance of the upper airway. I hypothesized that in cough/respiratory sounds, formant may carry resonances of the entire respiratory tract. For instance, wheezing sounds, which originates due to vibrations in the bronchioles of the lung, may contribute higher frequency formants (resonance frequencies) in the cough sounds. The first four FF was included in wet/dry classification work, and five formants were used in the automatic segmentation work. Past studies in the speech and acoustic analysis have shown that these formants correspond to various acoustic features of airways [147, 148]. I computed the formant frequencies by peak picking the Linear Predictive Coding (LPC) spectrum of cough sounds. For this work, I used the 14th order LPC model with the parameters determined via the Levinson-Durbin recursive procedure.

A1.6. Pitch

In speech analysis, pitch is defined as the fundamental frequency of the vocal cord. Several algorithms have been proposed in the literature to estimate the pitch of a voiced acoustic signal. In this study, I used the classical method of 'autocorrelation with center clipping' [150] to compute the pitch of a cough sub-segment.

A1.7. Log energy

The log energy for every sub-segment was computed using eq. A-15.

$$\mathrm{E}_k = 10\log 10\left(\varepsilon + \frac{1}{N}\sum_{n=1}^{N} \hat{s}_k(n)^2 \right) \qquad \text{(A-15)}$$

In (15) ε in (%) is an arbitrarily small positive constant added to prevent any inadvertent computation of the logarithm of 0.

A1.8. Kurtosis

Kurtosis is a measure of the "peakedness" associated with a probability distribution of sub-segment $\hat{s}_k[n]$, computed using (A-16). μ and σ are the mean and standard deviation of the segment $\hat{s}_k[n]$ respectively.

$$\lambda_k = \left(\frac{1}{N}\sum_{n=1}^{N} \left(\frac{\hat{s}_k[n] - \mu}{\sigma} \right)^4 \right) - 3 \qquad \text{(A-16)}$$

A1.9. Bispectrum score (BSS)

The third order spectrum of the signal is known as the bispectrum. Unlike the power spectrum (second order statistics) based on the autocorrelation, bispectrum preserves Fourier phase information. The bispectrum can be estimated via estimating the third order cumulant and then taking a 2D-Fourier transform. The third order cumulant $C(\tau_1, \tau_2)$ was estimated using (A-17) as defined in [151]. By applying a bispectrum window function (minimum bispectrum-bias supremum window described in [152]) to the cumulant estimate, a windowed cumulant function $C_k(\tau_1, \tau_2)$was obtained.

$$C_k(\tau_1,\tau_2)=\frac{1}{N}\sum_{n=1}^{N-1}\hat{s}_k(n)\hat{s}_k(n+\tau_1)\hat{s}_k(n+\tau_2),\qquad |\tau_1|\le Q,\ |\tau_2|\le Q \tag{A-17}$$

In (17) Q is the length of the 3rd order correlation lags considered. The bispectrum $B_k(\omega_1,\omega_2)$ of the sub-segment $\hat{s}_k[n]$ was estimated using (A-18). FFT length is set at 512 points.

$$B_k(\omega_1,\omega_2)=\sum_{\tau_1=-\infty}^{\tau_1=+\infty}\sum_{\tau_2=-\infty}^{\tau_2=+\infty}C_k(\tau_1,\tau_2)e^{-j(\tau_1\omega_1+\tau_2\omega_2)} \tag{A-18}$$

In the frequency domain, a quantity $P_k(\omega;\varphi,\rho)$ can be defined for the data sub-segment $\hat{s}_k[n]$ such that:

$$P_k(\omega;\varphi,\rho)=B_k(\omega,\varphi\omega+\rho) \tag{A-19}$$

describing a one-dimensional slice inclined to the ω_1-axis at an angle $tan^{-1}\varphi$ and shifted from the origin along the ω_2-axis by the amount ρ,(-π<ρ<π). For this work I set φ=1 and ρ=0 so that the slice of the bispectrum considered is inclined to the ω_1-axis by 45^o and passes through the origin. Then Bispectrum Score ξ is computed using (A-20). In (A-20) I used ω_1 = 90 Hz, ω_2 = 5 kHz, ω_3 = 6 kHz and ω_4 = 10.5 kHz.

$$\xi=\frac{\int_{\omega_1}^{\omega_2}P(\omega)}{\int_{\omega 3}^{\omega_4}P(\omega)} \tag{A-20}$$

A.2. Performance parameters

Definition of the statistical measures used to evaluate the performance of the algorithm.

$$\text{Sensitivit y}=TP/(TP+FN) \tag{A-21}$$

$$\text{Specificit y}=TN/(FP+TN) \tag{A-22}$$

$$\text{Accuracy }=(TP+TN)/(TP+FP+TN+FN) \tag{A-23}$$

$$\text{PPV}=TP/(TP+FP) \tag{A-24}$$

$$\text{NPV}=TN/(TN+FN) \tag{A-25}$$

where TP – True Positive, FP – False Positive, TN – True Negative, FN – False Negative, PPV – Positive Prediction Value, NPV – Negative Prediction Value.

In this work, Kappa statistic measures the agreement between the proposed algorithm and the trained scorer. Below are the guidelines for interpreting the Kappa values.

Table A.1: The interpretation of Kappa values.

Kappa	Interpretation
< 0	less than chance agreement
0.01 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement

A.3. Recording devices specification

In this study, I used Rode NT3 microphones incorporated with pre-amplifier and A/D converter MobilePre USB, M-Audio®, California, USA. The specification of the microphones is illustrated in Table A3.1 while the polar pattern and frequency response are shown in Fig A3.1 and Fig A3.2, respectively. The frequency response of MobilePre USB M-Audio is shown in Fig A3.3.

Table A.2: The specification of microphone RODE NT3.

Acoustic principle	Pressure gradient
Directional pattern	Hypercardioid
Sensitivity	-39 dB re 1Volt/Pascal(12 mV @ 94 dB SPL) +/- 2 dB @ 1 kHz
Output impedance	200Ω
Output connection	3 pin XLR, balanced output between Pin 2 (+), Pin 3 (-) and Pin 1 (ground)
Power	9V battery or phantom power 12V
SPL/Noise	Max sound pressure level (SPL): 140 dB, Self-noise: 16 dB (A)

A.4. Graphical user interface for wet/dry cough classification

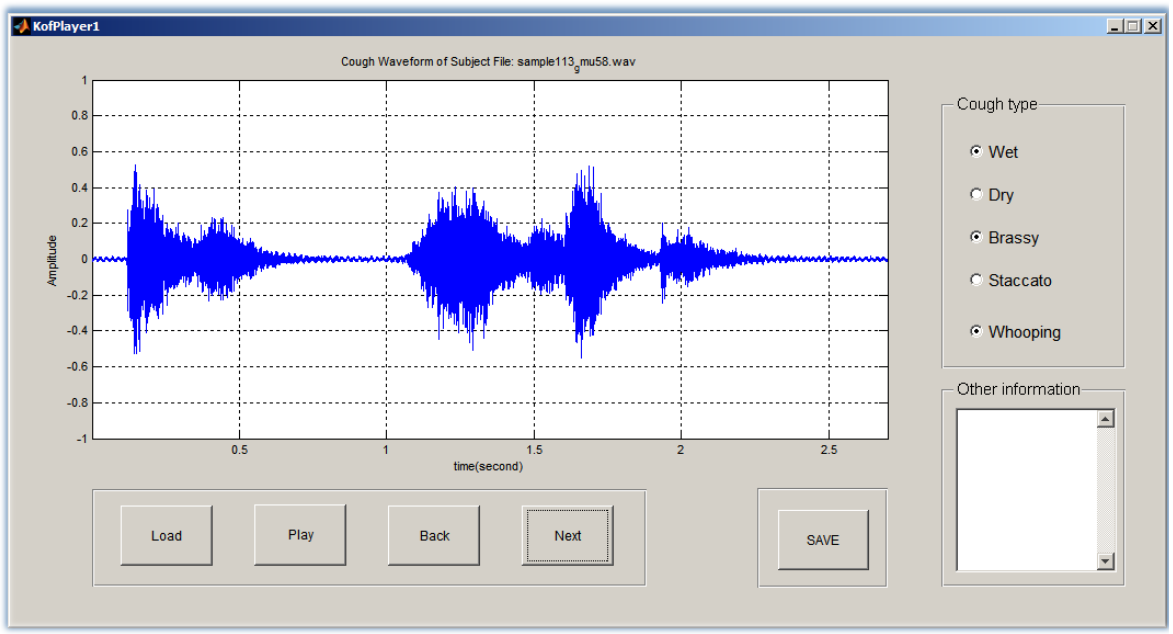

Figure A.1: Graphical user interface application (GUI) used by paediatricians for classifying coughs into different categories.

A.5. Feature vectors statistics from cough sound and non-cough sounds

The feature vectors *F* for automatic cough segmentation contains 22 elements: 13 Mel-frequency cepstral coefficients, 5 Formant Frequency, and 1 coefficient of each Log Energy, Zero crossing rates, Shannon entropy, and Non-Gaussianity score. To discover the characteristic of these features, I calculated the probability density function (*pdf*) of a specific sound in the cough (CG) and non-cough (NG) classes. The NG represent a wide range of non-cough sounds; hence, in this work, I chose the most dominant sounds, such as cry (CY), vocalization (abbreviated as VC, e.g., speech, typical baby voices), and appliances sounds (abbreviated as AS, e.g., sound from the door bank, trolley, bed).

I illustrate the smoothed *pdf* of each feature of *F* in Figures A.4 through A.6. All values in the *pdf* had been normalized. As can be seen from these figures, the distribution of features (MFCC, FF, CR SH, and NGS) between cough and other sounds are overlapping. However, each component of the features has a unique distribution.

Figure A.4 shows the *pdf* from randomly selected elements of MFCC (M(1), M(4), M(9), and M(11)). From Figures A.4(A) and A.4(B), it can be seen that the *pdf* of M(1) in CG and AS classes have a different mean (μ = 0.31 to 0.18). The *pdf* of M(9) in Figure A.4(C) shows that CG has lower mean than CY (μ = -0.01 to μ = -0.13, respectively). The complete statistical information of MFCC is shown in Table A.3.

Figure A.5 exhibits the *pdf* of formant frequencies. The statistical distribution (mean, standard deviation, skewness, and kurtosis) of formant frequencies of F(1), F(3), and F(5) between cough (CG) and appliance sounds (AS) are distinguishable. The distribution of F(2) in CG, CY, VC, and AS seem similar; however, CG has the lowest mean (-0.06). Moreover, the distribution of F(4) in TS has the lowest mean among the classes (-0.003).

Figure A.6 illustrates the NGS index of CG, CY, VC, and AS. The NGS has the potential to discriminate CG from VC and CY (Figure A.6(D)). Similarly, from Figure A.6(B) it can be seen that ZCR can be used to discriminate CG from AS. I show the mean and standard deviation of formant frequency, Shannon entropy, zero crossing rates, and non-Gaussianity score in Tables A.4 and A.5.

The *pdf* of the features shows that there are no dominant feature which can be used alone as an input for TDNN to classify CG/NG class. Hence, to obtain the maximum benefit of each component of the features, I combined them and used *d* (*d* = 5) successive of features vector as the input of TDNN to classify CG/NG class.

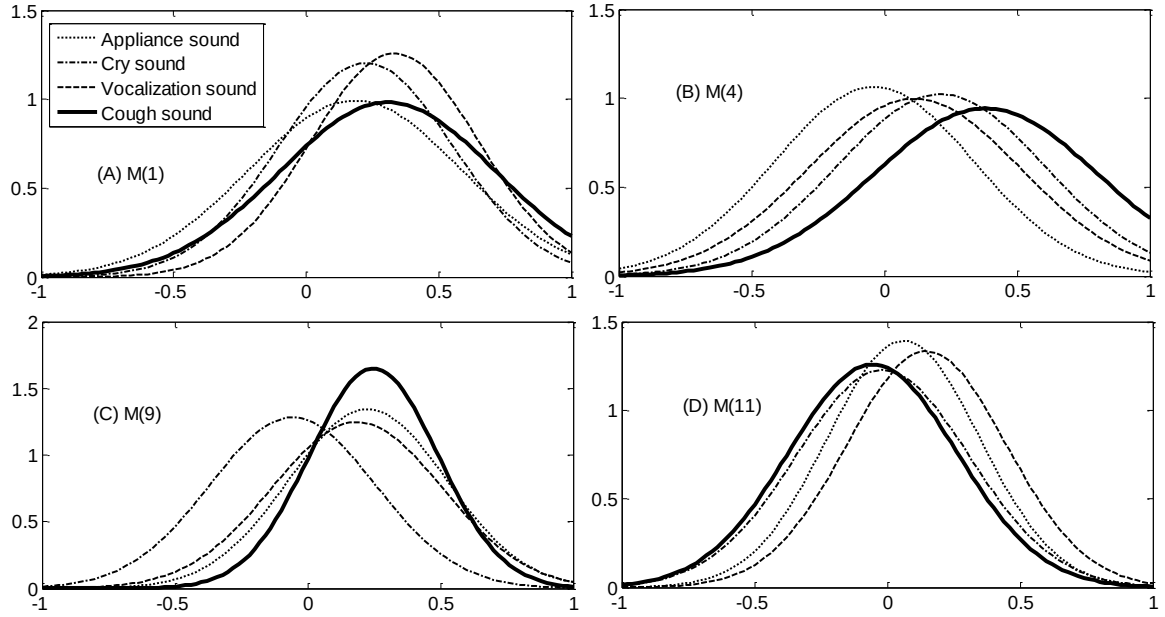

Figure A.2: The probability density function (*pdf*) of randomly selected Mel-frequency cepstral coefficient (smoothed for display purposes). Although they are overlapping, the coefficient M(4) can be used to differentiate cough and appliance sound, the coefficient M(9) differentiate between cough to cry, and M(11) to differentiate between cough and vocalization.

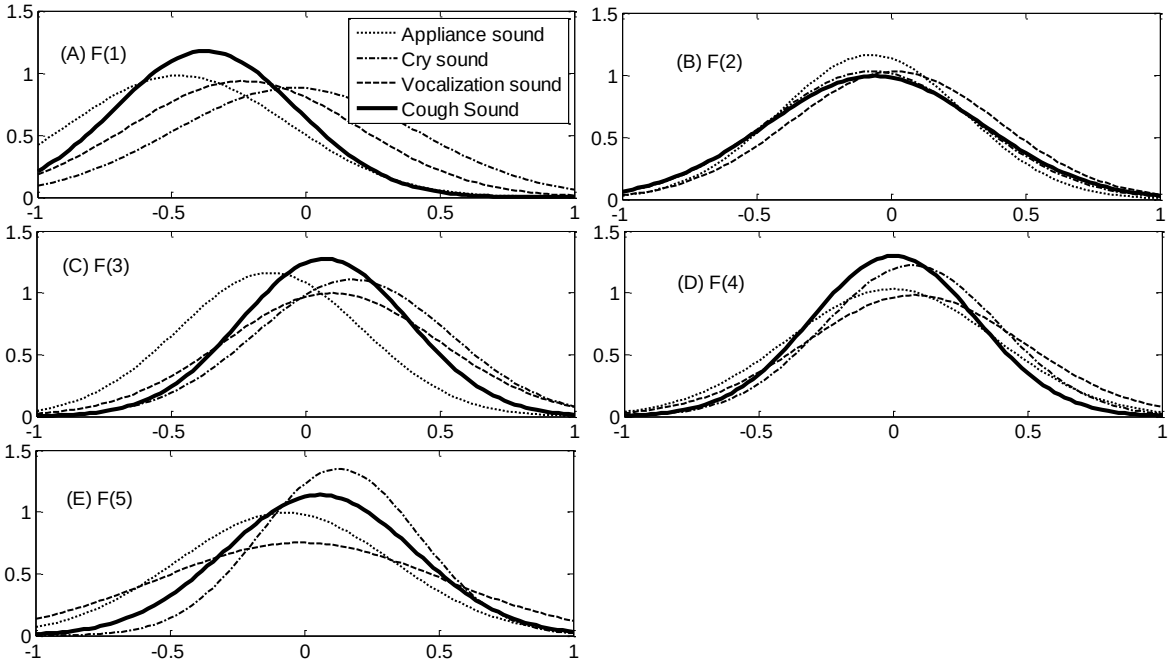

Figure A.3: The *pdf* of five first formant frequencies (F(1) – F(5)). Even though the distribution of the formant frequencies are overlapping, they have different mean, skewness, and kurtosis, especially for F(1), F(3), and F(5).

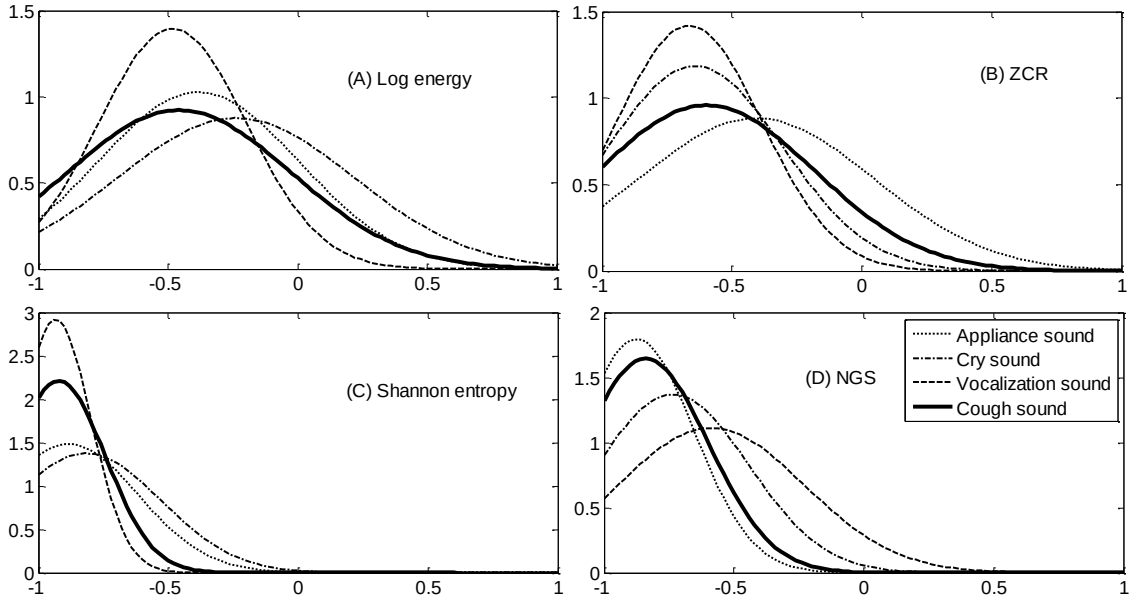

Figure A.4: The *pdf* of the energy, zero crossing rate (ZCR), Shannon entropy, and non-Gaussianity score (NGS). The NGS can be used to differentiate between cough with vocalization and cry, while ZCR differentiates between cough and appliance sound. The profile of log energy and Shannon entropy between sounds are similar, but they have different skewness and kurtosis.

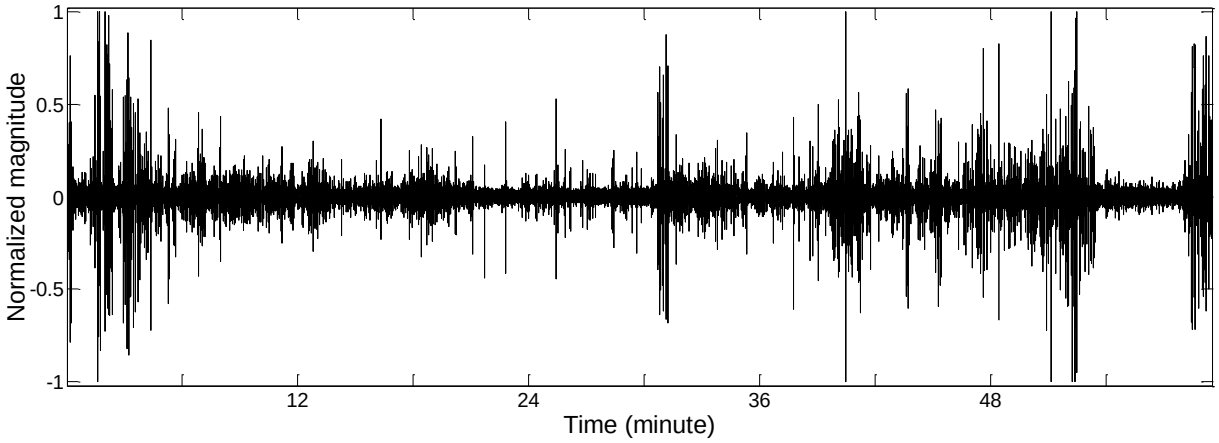

Figure A.5: Illustration of cough sound recording. As well as coughs, the recording also contains an abundance of other sounds from children, such as crying, grunting, and vocalizations.

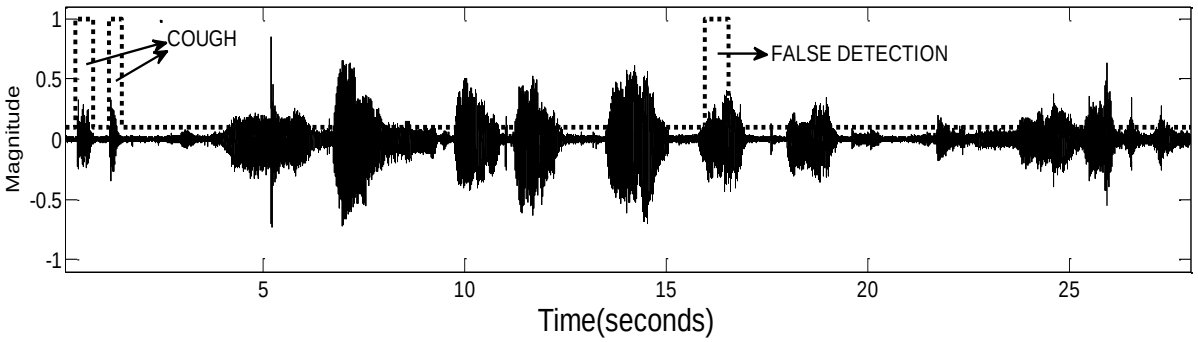

Figure A.6: A portion of an enlarged cough sound recording along with the segmentation results (indicated by the dashed line). It shows 30 second length recordings containing coughs, crying, and grunting.

Table A.3: Mel-frequency cepstral coefficients (MFCCs) statistics from cough (CG), Vocalization (VC), Cry (CY) and Appliances sounds (AS). M(1) – M(12) represent the MFCCs, μ = mean and σ = standard deviation.

Class	M(1)		M(2)		M(3)		M(4)		M(5)		M(6)	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
AS	0.18	0.40	0.18	0.42	-0.01	0.36	-0.04	0.38	0.23	0.30	0.17	0.35
CY	0.23	0.33	0.36	0.35	0.08	0.34	0.21	0.39	-0.06	0.31	0.09	0.33
VC	0.33	0.32	0.46	0.28	-0.13	0.31	0.11	0.40	0.19	0.32	0.33	0.33
CG	0.31	0.41	0.23	0.38	0.33	0.26	0.38	0.42	0.24	0.24	0.27	0.37

Class	M(7)		M(8)		M(9)		M(10)		M(11)		M(12)	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
AS	0.04	0.37	0.09	0.28	-0.13	0.31	0.18	0.30	0.06	0.29	0.04	0.36
CY	-0.03	0.32	0.08	0.30	-0.12	0.31	0.06	0.33	-0.02	0.33	0.16	0.32
VC	-0.29	0.28	-0.07	0.32	0.02	0.30	-0.23	0.28	0.15	0.30	-0.08	0.30
CG	-0.02	0.33	-0.14	0.29	-0.01	0.30	-0.10	0.34	-0.05	0.32	0.03	0.35

Table A.4: Formant frequency (F(1) – F(5)) statistics from cough (CG), Vocalization (VC), Cry (CY) and Appliances sounds (AS).

Class	F(1)		F(2)		F(3)		F(4)		F(5)	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
AS	-0.47	0.41	-0.07	0.34	-0.13	0.34	-0.003	0.39	0.07	0.39
CY	-0.04	0.45	-0.07	0.39	0.18	0.36	0.07	0.33	0.12	0.30
VA	-0.23	0.43	0.01	0.39	0.10	0.40	0.08	0.41	0.01	0.53
CG	-0.37	0.33	-0.06	0.40	0.07	0.31	0.003	0.30	0.06	0.35

Table A.5: Log energy, zero crossing rates (ZCR), Shannon entropy, and non-Gaussianity score (NGS) statistics from cough (CG), vocalization (VC), cry (CY) and appliances sounds (AS).

Class	Log Energy		ZCR		Shannon Ent.		NGS	
	μ	σ	μ	σ	μ	σ	μ	σ
AS	-0.38	0.39	-0.41	0.45	-0.88	0.27	-0.87	0.22
CY	-0.24	0.46	-0.64	0.34	-0.82	0.29	-0.73	0.29
VC	-0.48	0.29	-0.66	0.28	-0.93	0.14	-0.59	0.36
CG	-0.46	0.43	-0.60	0.42	-0.92	0.18	-0.84	0.24

Table A.6: Automatic segmentation results of each subject in Prospective Study Dataset (PSD). CG = cough, NC = Non-cough, TP, True Positive, FN = False Negative, TN = True Negative, FP = False Positive, Sen = Sensitivity, Spe = Specificity, Acc = Accuracy, K = Kappa, PPV = Positive Predictive Value, NPV = Negative Predictive Value.

Subject ID	CG	NC	TP	FN	TN	FP	Sen (%)	Spe (%)	Acc (%)	κ	PPV (%)	NPV (%)
1	275	2553	251	24	2499	54	91.3	97.9	97.2	0.9	82.3	99.0
2	23	2311	19	4	2254	57	82.6	97.5	97.4	0.4	25.0	99.8
3	57	2226	45	12	2212	14	78.9	99.4	98.9	0.8	76.3	99.5
4	100	4555	97	3	4389	166	97.0	96.4	96.4	0.5	36.9	99.9
5	96	5039	94	2	4887	152	97.9	97.0	97.0	0.5	38.2	100.0
6	270	2413	260	10	2320	93	96.3	96.1	96.2	0.8	73.7	99.6
7	172	3598	140	32	3425	173	81.4	95.2	94.6	0.6	44.7	99.1
8	112	2542	105	7	2489	53	93.8	97.9	97.7	0.8	66.5	99.7
9	71	4315	68	3	4231	84	95.8	98.1	98.0	0.6	44.7	99.9
10	66	2679	61	5	2610	69	92.4	97.4	97.3	0.6	46.9	99.8
11	51	2857	50	1	2785	72	98.0	97.5	97.5	0.6	41.0	100.0
12	29	1526	29	0	1502	24	100.0	98.4	98.5	0.7	54.7	100.0
13	20	1376	17	3	1361	15	85.0	98.9	98.7	0.6	53.1	99.8
14	87	2154	87	0	2108	46	100.0	97.9	97.9	0.8	65.4	100.0