34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Automated algorithm for Wet/Dry cough sounds classification

Swarnkar V, Abeyratne U R, *Senior Member IEEE*, Yusuf A. Amrulloh, *member IEEE*, Anne Chang

Abstract— **Cough is the most common symptom of several respiratory diseases. It is a defense mechanism of the body to clear the respiratory tract from foreign materials inhaled accidentally or produced internally by infections. The identification of wet and dry cough is an important clinical finding, aiding in the differential diagnosis. Wet coughs are more likely to be associated with bacterial infections. At present, the wet/dry decision is based on the subjective judgment of a physician, during a typical consultation session. It is not available for long term monitoring or in the assessment of treatment efficacy. In this paper we address these issues and develop fully automated technology to classify cough into 'Wet' and 'Dry' categories. We propose novel features and a Logistic regression-based model for the classification of coughs into wet/dry classes. The performance of the method was evaluated on a clinical database of pediatric and adult coughs recorded using a bed-side non-contact microphone. The sensitivity and specificity of the classification were obtained as 79±9% and 72.7±8.7% respectively. These indicate the potential of the method as a useful clinical tool for cough monitoring, especially at home settings.**

I. Introduction

Cough is a natural protective mechanism that helps clearing the secretions from the respiratory tract and prevents entering of noxious particles into the respiratory system. It is generally defined as the sudden expulsion of air accompanied with a typical sound [1]. The prevalence of cough in communities in Europe and USA varies between 9 – 33%[2]. The situation is far worse in the developing world.

Cough can be classified into the two categories 'Wet Cough' and 'Dry Cough' depending on their 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. This is essentially a subjective process.

Medically there are different reasons for the wet and dry cough and their identification aids in the differential diagnosis of diseases such as pneumonia and bronchiolitis. Often, the dry-wet classification is used in epidemiological studies [3, 4] and clinical research [5, 6]. In children wet cough is generally associated with the lower respiratory tract infections [6]. Diseases such as bronchiolitis, allergies, sinusitis can cause dry cough.

Cough is often present as an earliest symptom in almost all of the respiratory diseases. It can be a useful in developing screening tools for some respiratory diseases. Even though cough is common in respiratory diseases and considered an importance clinical symptom, there is no objective gold standard to assess it. Manual assessment of dry and wetness of the cough sounds is the reference method used by clinicians around the globe [5].

Researchers have rarely attempted to develop technology for the automated, objective classification of cough into dry-wet categories. To the best of our knowledge, only two prior works exist in this area [7, 8]. Murata et al [7] analyzed cough sound frequencies to discriminate between wet and dry cough. Chatrzarrin et al [8] proposed peaks of the energy envelop and spectral features of the cough sounds to differentiate between wet and dry cough.

These studies opened up a new branch of research in respiratory sound analysis. However they have been 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 analyzed was limited (30 cough samples from 10 subjects in [7] and total of 16 coughs in [8]) making the interpretation of the results difficult. All the existing work used cough sounds from adult subjects. In addition to these, characterizations of the cough sounds were based on duration, magnitude and frequency features.

Production of cough sound is a complex physiological process involving several anatomical structures in the lower and upper respiratory system. Its characteristic features vary significantly with the individual differences and depends heavily on respiratory conditions [1]. Intensity and duration dependent methods will not be sufficient to capture the rich information hidden in cough sounds.

In this paper we propose an automated classification model to categorize cough sounds into wet and dry groups. Method uses 1st, 2nd and 3rd order statistical features (eg. formant frequencies, mel-cepstrum, non-Gaussianity, and bispectrum etc.) of the cough sounds. Model is trained and tested on a comprehensive database of 178 coughs from 46 subjects (23 male, 23 female) with age range of 1 month to 15 years. The subjects have a range of respiratory illnesses such as asthma, pneumonia, bronchitis and rhinopharyngitis.

II. Method

A. Recording environment

The clinical data acquisition environment for this work is Respiratory Medicine Unit of the Sardjito Hospital, Gadjah Mada University, Indonesia. Table 1 lists the inclusion and exclusion criteria. All patients fulfilling the inclusion criteria were approached. An informed consent was made using form approved by the human ethics committes of Gadjah Mada University and The University of Queensland. Patients

This work has been partially supported by a Grand Challenges in Global Health Exploration Grant to Udantha from The Bill & Mellinda Gates Foundation, USA. Vinayak Swarnkar, Udantha Abeyratne and Yusuf Amrulloh are with the School of ITEE, The University of Queensland, 4072, Australia. Anne Chang is with Royal Children's Hospital Brisbane, Australia. Corresponding author: udantha@itee.uq.edu.au.

TABLE I
INCLUSION AND EXCLUSION CRITERIA USED IN THE STUDY

Inclusion Criteria	Exclusion Criteria
- Patients with symptoms of chest infection : At least 2 of - Cough - Sputum - Increased breathlessness - Temperature >37.5° - Consent	- Advanced disease where recovery is not expected eg terminal lung cancer - Droplet precautions - NIV required - No Consent

were recruited within first 12 hours of their admission. After the initial medical assessment sound recordings were made for next 4-6 hours in the natural environment of the respiratory ward.

The cough sounds were acquired with a high fidelity, computerized data acquisition system. A matched pair of low-noise microphones having a hypercardiod beam pattern (Model NT3, RODE, Sydney, Australia) were used to capture the sound signals. The nominal distance from the microphone to the mouth of the patient was 50 cm, but could vary from 40 cm to 70 cm due to patient movements. A professional quality pre-amplifier and A/D converter unit (Model Mobile-Pre USB, M-Audio, California, USA) was used for sound signal acquisition. We used a sampling rate of 44.1k samples with a 16 bit resolution (CD-quality recording). The system had a 3 dB bandwidth of 20800 Hz at the sampling rate used.

B. Formation of feature vector and Wet/Dry cough classification model

There is no widely accepted method for automatic identification of cough events. Manual identification of cough events from long sound recording is still considered as the best method. In this paper we followed this standard. After the manual scoring we followed the following steps to compute mathematical features from the cough event data.

[1]. Let x[k] denotes the k^{th} sample of the discrete time sound signal. Filter x[k] using a digital high pass filter to get y[k].

[2]. Divide y[k] into 'n' equal size non-overlapping sub-segments. Let $y^i[k]$ represents the i^{th} sub-segment of y[k], where i = 1,2,3,….,n.

[3]. Compute the following features (see Section C for details) for each of the 'n' sub-segments in y[k]: Bispectrum Score (BGS), Non-gaussianity score (NGS), formant frequencies (FF), log energy (LogE), zero crossing (ZCR), kurtosis (Kurt), and mel-frequency cepstral coefficients (MFCC).

[4]. For each y[k] form a feature vector F_k containing N elements where N consists of:
(12 × n from MFCC) +(4 × n from Formant frequency) +(5 × n from NGS, LogE, Zcr, Kurt and Bispectrum). By setting n=3 we get an F_k with N=63 features for each cough event.

[5]. For Wet and Dry cough classification we used Logistic Regression (LR) statistical model. It is a generalized linear model, which uses independent several 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 LR function to estimate the probability Y=1 (i.e cough event belong to category of 'Wet Cough') given the independent variables (i.e feature set) as follows:

$$Prob\big(Y = 1|_{x1,x2,x3,\dots xn}\big) = \frac{e^z}{e^z + 1} \quad (1)$$

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (2)$$

Features were then selected to include only the best independent variables (variables with low 'p' value) that facilitate the classification, in the final model. The final model is then used to estimate the probability and each cough event is classified as belonging to either of the two categories using a probability threshold.

C. Feature computation

(i) Bispectrum Score (BGS) – The 3rd order spectrum of the signal is known as the bispectrum. Unlike the power spectrum (2nd order statistics) based on the autocorrelation, bispectrum preserves Fourier phase information. The bispectrum can be estimated via estimating the 3rd order cumulant and then taking a 2D-Fourier transform. The 3rd order cumulant $C(\tau_1,\tau_2)$ was estimated using (3) as defined in [9]. By applying a bispectrum window function (minimum bispectrum-bias supremum window described in [10]) to the cumulant estimate, windowed cumulant function $C^i_w(\tau_1,\tau_2)$was obtained.

$$C^i(\tau_1,\tau_2) = \frac{1}{L}\sum_{k=0}^{L-1} y^i(k)y^i(k+\tau_1)y^i(k+\tau_2), \quad |\tau_1| \le Q, |\tau_2| \le Q \quad (3)$$

In (3) Q is the length of the 3rd order correlation lags considered. The bispectrum $B^i(\omega_1,\omega_2)$ of the segment $y^i[k]$ was estimated using (4). We used FFT length of 512 points.

$$B^i(\omega_1,\omega_2) = \sum_{\tau_1=-\infty}^{\tau_1=+\infty}\sum_{\tau_2=-\infty}^{\tau_2=+\infty} C^{yi}(\tau_1,\tau_2)\, e^{-j(\tau_1\omega_1+\tau_2\omega_2)} \quad (4)$$

In the frequency domain, a quantity $P^i(\omega;\phi,\rho)$ can be defined for the data segment $y^i[k]$ such that

$$P^i(\omega;\varphi,\rho) = B^i(\omega,\varphi\omega+\rho) \quad (5)$$

describing a one-dimensional slice inclined to the ω_1-axis at an angle $tan^{-1}\phi$ and shifted from the origin along the ω_2-axis by the amount ρ, $(-\pi < \rho < \pi)$ [9]. For this work we set ϕ=1 and ρ=0 so that the slice of the bispectrum considered is inclined to the ω_1-axis by 45° and passes through the origin. Then Bispectrum Score (BSG) is computed using (6). In (6) we used ω_1 = 90hz, ω_2 = 5khz, ω_3 = 6khz and ω_4 = 10.5khz.

$$BSG = \frac{\int_{\omega_1}^{\omega_2} P(\omega)}{\int_{\omega_3}^{\omega_4} P(\omega)} \quad (6)$$

(ii) Non-Gaussianity Score (NGS) – NGS gives the measure of non-gaussianity of a given segment of data. The normal

probability plot can be utilized to obtain a visual measure of the gaussianity of a set of data. The NGS of the data segment $y^i[k]$ can be calculated using (7). Note that in (7), p and q represents the normal probability plot of the reference normal data and the analyzed data, respectively, with j ranging from the values 1 to N.

$$NGS = 1 - \left(\frac{\sum_{j=1}^{N}(q[j]-p)^2}{\sum_{j=1}^{N}(q[j]-\bar{q})^2}\right) \quad (7)$$

(iii) Formants frequencies – In human voice analysis formants are referred as the resonance of the human vocal tract. In cough analysis, it is reasonable to expect that the resonances of the overall airway that contribute to the generation of a cough sound will be represented in the formant structure; mucus can change acoustic properties of airways. We included the 1st four formant frequencies (F1, F2, F3, F4) in our feature set. Past studies in the speech and acoustic analysis have shown that F1-F4 corresponds to various acoustic features of airway[11]. We computed the F1-F4 by peak picking the Linear Predictive Coding (LPC) spectrum of cough sounds. For this work we used 14th order LPC model with the parameters determined via the Levinson-Durbin recursive procedure [12].

(iv) Log Energy(LogE) – The log energy for every sub-segment was computed using eq. 8

$$LogE = 10log_{10}\left(\varepsilon + \frac{1}{N}\sum_{k=1}^{N}(y^i(k)^2)\right) \quad (8)$$

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

(v) Zero crossing (Zcr) – The number of zero crossings were counted for each 'n' segments.

(vi) Kurtosis (Kurt)– The kurtosis is a measure of the peakedness associated with a probability distribution of segment $y^i[k]$, computed using (9). μ and σ is the mean and stand deviation of the segment $y^i[k]$ respectively.

$$kurt = \frac{E(y^i[k]-\mu)^4}{\sigma^4} \quad (9)$$

(vii) Mel-frequency cepstral coefficients (MFCC) – MFCCs are commonly used in the speech analysis systems [13]. They represent the short term power spectrum of an acoustic signal based on a cosine transform of a log power spectrum on a non-linear mel-scale of frequency. We included the 12 MFCC coefficients in our feature set.

TABLE II
MEAN ± STD VALUES FOR SENSITIVITY, SPECIFICITY, POSITIVE PREDICTIVE VALUE (PPV), NEGATIVE PREDICTIVE VALUE (NPV), ACCURACY AND KAPPA, FOR 200 DESIGNED LR MODELS

200 Datasets	Using all the cough features		Using selected cough features	
	Training	Testing	Training	Testing
Sensitivity	100±0.1	74.8±9	95.93±2	79±9
Specificity	100±0	69.9±9.4	83.4±4	72.7±8.7
PPV	100±0.1	75.7±6.4	80.9±3.9	78.3±5.6
NPV	100±0	70±8	96.6±1.6	74.4±8.6
Accuracy	100±0.1	72.6±6.1	88.7±2.7	76.1±5.5
Kappa	1±0	0.45±0.12	0.77±0.05	0.52±0.1

III. RESULTS

A. Training and testing datasets

Total of 178 cough events from 46 subjects were used in this study. The male to female ratio of the 46 subjects was 1:1. The mean age of the subjects was 3 years and 3 month. The age range of the subjects varied from 1 month to 15 years and having diseases such as asthma, pneumonia, bronchitis, rhinopharyngitis etc.

The fourth author of the paper, a pediatrician with more than 20 years of experience in pediatric respiratory diseases, manually classified 178 cough events into Wet and Dry. The pediatrician was blinded to the actual diagnosis of the subjects. This manual classification was considered as the 'reference standard' against which results of automatic classification by designed LR model were compared.

Out of 178 cough events 82 were classified as Wet and 96 as Dry. We randomly partitioned the data set into non-overlapping training and testing sets. We used 70% (124 cough events) of the cough events for training and 30% (54 cough events) for testing the model. The model was trained on the training set of each partition and was tested on the particular testing set. Each partition thus gave us a performance indication of our method. To validate the model, we generated 200 such randomized partitions and evaluated the average performance over them.

B. Classification results

The mean sensitivity and specificity for Wet/Dry classification using LR-model was 74.8±9% and 69.9±9.4% respectively for testing datasets, when all the cough features were used to train the model. Mean sensitivity and specificity values jumped to 79±9% and 72.7±8.7% when only selected cough features were used. In all 22 features were selected out of 63 after the feature optimization. The p-value of 0.4 was used for feature selection. The selected features were 1 each from BSG, LogE and Kurt; 2 from NGS; 3 from ZCR; 5 from formant frequency; and 9 from MFCC. Table 2 shows the mean sensitivity, specificity, accuracy and kappa results for training and testing datasets. The kappa agreement (widely used statistic in situations where the agreement between two techniques is compared) between the LR-model and reference method was 0.52±0.1. Figure 1 show the histogram plots for the sensitivity and specificity obtained using 200 training and testing datasets. Table 3 shows the contingency table for the best LR-model among 200 and Fig.2 shows the model probability output (probability that the given cough is wet) and targeted

wet/dry cough class. It has the sensitivity of 90%, specificity of 80% and a high kappa agreement of 0.71.

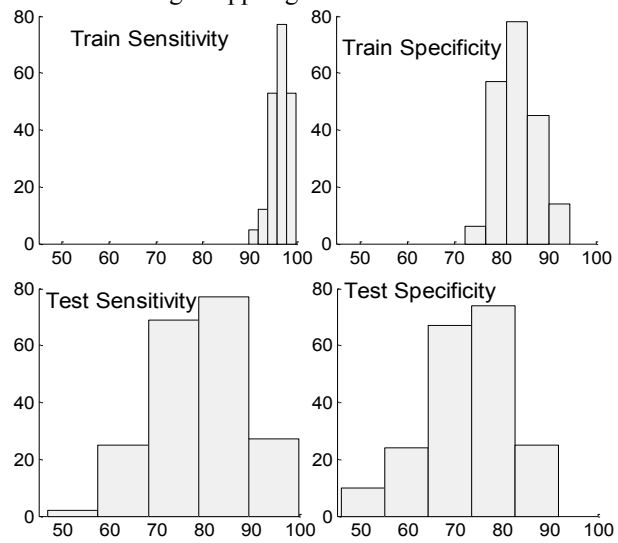

Fig 1. Histograms of sensitivity and specificities for 200 training and testing datasets. Only selected features were used for LR model designing.

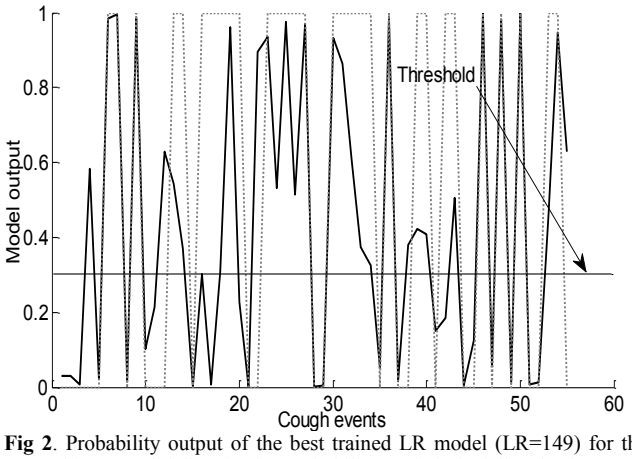

Fig 2. Probability output of the best trained LR model (LR=149) for the test subjects. Solid line is the probability and dashed line is the target wet/dry cough sound class. Threshold is the probability threshold applied to categorize cough into wet or dry. Note that the threshold was set at 0.3 instead of 0.5 normally used in LR modeling. This value was obtained after optimizing the results to get minimum mean training sensitivity of >90%, which can be seen in Fig.1

TABLE III
CONTINGENCY TABLE FOR BEST LR MODEL (LR=149)

		Test Method		
		W	**nW**	
Reference Standard	**W**	27	3	**30**
	nW	5	20	**25**
		32	**23**	**55**

IV. CONCLUSION

In this paper we presented an automated algorithm to classify Wet and Dry cough using mathematical features extracted from the cough sound. Our method can classify Wet and Dry coughs with high sensitivity (79%) and specificity (72.7%) and with a good agreement (kappa=0.52) with the expert human scorer. Proposed method carries the potential to develop as a useful clinical tool for long term cough monitoring, in the assessment of treatment efficacy or in characterizing the lower respiratory tract infections. This is the first known method for Dry/Wet classification, presented with complete training and testing results on significantly large cough samples (178 cough samples from 46 subjects). It is also the first effort to automate the Wet/Dry classification in pediatric population with range of respiratory infectious diseases. In the future, work will focus on increasing the data size and statistically analyzing the selected features by considering the impact of inclusion and exclusion of a certain feature.

ACKNOWLEDGMENT

Authors would like to thank Rina Triasih and Amelia Setyati, Sardjito Hospital, Yogyakarta, Indonesia for their assistance with cough data collection and clinical diagnosis as used in this paper.

REFERENCES

[1] J. Korpá, J. Sadloová, and M. Vrabec, "Analysis of the cough sound: an overview," *Pulmonary Pharmacology,* vol. 9, pp. 261-268, 1996.

[2] K. F. Chung and I. D. Pavord, "Prevalence, pathogenesis, and causes of chronic cough," *The Lancet,* vol. 371, pp. 1364-1374, 2008.

[3] M. E. Soto Quiros, M. Soto Martinez, and L. Å. Hanson, "Epidemiological studies of the very high prevalence of asthma and related symptoms among school children in Costa Rica from 1989 to 1998," *Pediatric allergy and immunology,* vol. 13, pp. 342-349, 2002.

[4] J. D. Spengler, J. J. K. Jaakkola, H. Parise, B. A. Katsnelson, L. I. Privalova, and A. A. Kosheleva, "Housing characteristics and children's respiratory health in the Russian Federation," *Journal Information,* vol. 94, 2004.

[5] A. B. Chang, J. T. Gaffney, M. M. Eastburn, J. Faoagali, N. C. Cox, and I. B. Masters, "Cough quality in children: a comparison of subjective vs. bronchoscopic findings," *Respir Res,* vol. 6, 2005.

[6] 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, pp. e364-e369, 2012.

[7] A. Murata, Y. Taniguchi, Y. Hashimoto, K. Y., Y. Takasaki, and S. Kudoh, "Discrimination of productive and non-productive cough by sound analysis," *Internal medicine,* vol. 37, pp. 732-735, 1998.

[8] H. Chatrzarrin, A. Arcelus, R. Goubran, and F. Knoefel, "Feature extraction for the differentiation of dry and wet cough sounds," 2011, pp. 162-166.

[9] U. Abeyratne, "Blind reconstruction of non-minimum-phase systems from 1-D oblique slices of bispectrum," 1999, pp. 253-264.

[10] 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.

[11] A. K. Ng, T. S. Koh, E. Baey, T. H. Lee, U. R. Abeyratne, and K. Puvanendran, "Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea?," *Sleep medicine,* vol. 9, pp. 894-898, 2008.

[12] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-time signal processing* vol. 1999: Prentice hall Englewood Cliffs, NJ:, 1989.

[13] F. Zheng, G. Zhang, and Z. Song, "Comparison of different implementations of MFCC," *Journal of Computer Science and Technology,* vol. 16, pp. 582-589, 2001.