



HHS Public Access

Author manuscript

IEEE J Biomed Health Inform. Author manuscript; available in PMC 2016 September 13.

Published in final edited form as:

IEEE J Biomed Health Inform. 2016 May ; 20(3): 963–969. doi:10.1109/JBHI.2015.2412880.

Classification of voluntary cough airflow patterns for prediction of abnormal spirometry

Jeffrey S. Reynolds,

Health Effects Laboratory Division, CDC/NIOSH, Morgantown, WV, USA

W. Travis Goldsmith,

Health Effects Laboratory Division, CDC/NIOSH, Morgantown, WV, USA

Jeremy B. Day, Ayman A. Abaza,

West Virginia High Technology Consortium Foundation, Fairmont, WV, USA

Ahmed M. Mahmoud,

Department of Systems and Biomedical Engineering, Cairo University, Giza, Egypt

Ali A. Afshari,

Health Effects Laboratory Division, CDC/NIOSH, Morgantown, WV, USA

Jacob B. Barkley,

Davis Memorial Hospital, Elkins, WV, USA

E. Lee Petsonk,

Division of Respiratory Disease Studies, CDC/NIOSH, Morgantown, WV, USA

Michael L. Kashon, and

Health Effects Laboratory Division, CDC/NIOSH, Morgantown, WV, USA

David G. Frazer

Health Effects Laboratory Division, CDC/NIOSH, Morgantown, WV, USA

Abstract

Measurement of partial expiratory flow-volume curves has become an important technique in diagnosing lung disease, particularly in children and in the elderly. The objective of this study was to investigate the feasibility of predicting abnormal spirometry using the partial flow-volume curve generated during a voluntary cough. Here, abnormal spirometry is defined as less than the lower limit of normal (LLN) predicted by standard reference equations [1].

Cough airflow signals of 107 subjects (56 male, 51 female) were previously collected [2] from patients performing spirometry in a pulmonary function clinic. A variety of features were extracted from the airflow signal. A support vector machine (SVM) classifier was developed to predict abnormal spirometry. Airflow signal features and SVM parameters were selected using a genetic algorithm. The ability of the classifier to distinguish between normal and abnormal spirometry based on cough flow was evaluated by comparing the classifiers decisions with the LLN for the given subject's spirometry, including forced expiratory volume in one second (FEV_1), forced vital capacity (FVC), and their ratio ($FEV_1/FVC\%$).

Findings indicated that it was possible to classify patients whose spirometry results were less than the LLN with an overall accuracy of 76% for FEV_1 , 65% for FVC , and 76% for the ratio $FEV_1/FVC\%$. Accuracies were determined by repeated double cross-validation [3].

This study demonstrates the potential of using airflow measured during voluntary coughing to identify test subjects with abnormal spirometry.

Index Terms

Cough air-flow spirometry support vector machine

I. Introduction

Spirometry testing remains one of the most widely used methods of evaluating the mechanical properties of the pulmonary system. Experimental evidence indicates that only a modest expiratory effort is required to reach an effort independent condition when executing a maximal expiratory flow-volume (MEFV) maneuver [4]. Since the MEFV relationship is effort independent, it not only gives information about the lung's mechanical properties, but also tends to be reproducible for a given test subject [5]. For a variety of reasons it is sometimes difficult to perform a complete MEFV maneuver. In these cases, partial expiratory flow-volume (PEFV) maneuvers have been shown to be useful in evaluating pulmonary function [6].

In the past, PEFV curves have also been used to study normal lung growth and development along with the effects of the disease process in infants and young children and those subjects whose lung function has been severely compromised by a lung disease [7], [8], [9]. Adler and Wohl [7] were able to reconstruct PEFV curves in infants by rapidly applying positive pressures around the infant body. Morgan, et al. [8] showed that children aged between 3–5 years were able to perform PEFV maneuvers, but they had to undergo a considerable amount of practice in order to perform an adequate and reproducible effort. Wall et al. [9] collected the PEFVs for 45 healthy children and 12 with generally mild lung disease caused by cystic fibrosis. They reported that within-subject, day-to-day, and among-subject variabilities in flow rates and lung volumes for healthy younger subjects that performed PEFVs were very similar to variabilities in older subjects who were able to produce full MEFVs.

In adults, PEFV curves have also been used to evaluate lung function in healthy and diseased subjects. Partial and complete expiratory flow-volume curves generated by asthmatic patients with spontaneous bronchospasm were acquired by Zamel et al. [10] while studying bronchodilatation. In addition, Barnes et al. [6] demonstrated that PEFV curves could be used to obtain dose-response curves after the inhalation of β_2 -adrenoceptor agonist aerosols in eight normal subjects.

It has been established that the partial flow-volume curve relationship generated during a voluntary cough, or cough partial expiratory flow-volume (CPEFV), is representative of a PEFV maneuver that is effort independent. It follows that the flow-volume relationship of a cough should also be repeatable when it is initiated from the same lung volume [11], [5].

Beardsmore et al. [12] extracted several flow parameters based on airflow measurements made during voluntary coughs, such as peak airflow and the ratio of MEFV-equivalent airflow to peak airflow. The same investigations examined differences in these parameters between a group of normal children and a similar group of asthmatic children. Other groups of investigators have developed methods for simultaneously recording both cough sound and airflow characteristics during a cough [13], [14]. Additionally, investigators in our laboratory [2] developed classification algorithms which used cough sound and flow data to identify subjects with abnormal lung function.

There are several benefits of using a voluntary cough to obtain an estimate of the PEFV curve. They include: (1) executing a voluntary cough is very familiar to test subjects and little apprehension is encountered in performing the maneuver; (2) minimal training is required for a subject to perform an acceptable cough and the procedure requires a nominal amount of time to perform; (3) in many cases, the cough maneuver is a feasible procedure for populations that cannot perform the MEFV maneuver (very young and old subjects and those who have debilitating pulmonary diseases); (4) the training time for personnel administering the testing procedure is minimal; and (5) cough airflow patterns are reproducible when performed from the same initial lung volume preceded by the same volume history [5].

In our previous work, a system was developed to record both the sound and the airflow produced during a voluntary cough [13]. Subsequently, we showed that information contained with the cough sound and airflow together allowed for classification of subjects diagnosed with either normal lung function or with obstructive lung disease [2]. In this work, the data collected in [2] was reanalyzed while considering the cough flow signal only and disregarding the cough sound data. The objective was to determine if the cough airflow, individually, contained sufficient information to accurately classify subjects with abnormal spirometry. Rather than classify based on a clinical diagnosis that contains a subjective component, the classifier in this work is based on the spirometric pulmonary function measurements.

A preliminary conference abstract of this work has been published [15].

II. Methods

A. Study Design

Cough airflow patterns were recorded for volunteer subjects who had spirometry testing performed at the pulmonary function laboratory at Ruby Memorial Hospital (Morgantown, WV). In this study, three common spirometric variables were utilized: forced expiratory volume in one second (FEV_1), forced vital capacity (FVC), and the ratio of FEV_1 to FVC expressed as a percentage ($FEV_1/FVC\%$). Spirometric results for FEV_1 , FVC , and $FEV_1/FVC\%$ were considered abnormal if they were less than the lower limit of normal calculated using reference equations from Hankinson, et al. [1]. Cough airflow characteristics, along with demographic data, were used to develop classifiers for predicting the abnormal spirometric results.

B. Cough Airflow Measurements

The system used in this research was originally designed to simultaneously measure airflow and sound pressure during a cough. For the purpose of this study, only the airflow measurements were examined. The system design and calibration protocol has previously been described in detail [13]. A cylindrical mouthpiece was attached to a 2.54 cm diameter metal tube connected through a flexible tube of the same diameter to a pneumotachograph (Model Fleisch No. 2, Fleisch, Epalinges, Switzerland). The pressure drop across the pneumotachograph, which is proportional to airflow, was monitored using a differential pressure transducer (Model 239, Setra Systems, Boxborough, MA). Airflow waveforms were filtered with a third order lowpass Butterworth filter with a cutoff frequency of 50 Hz. Cough airflow waveforms were analyzed and CPEFV curves were obtained by integrating the airflow curve with respect to time. The system was designed and calibrated to minimize airflow measurement errors [13].

C. Test Subjects

Participation was requested from patients whose physicians ordered pulmonary function testing at the Ruby Memorial Hospital laboratory. The study protocol was reviewed and approved by the Institutional Review Board of West Virginia University. After obtaining informed consent, investigators recorded voluntary coughs for 107 test subjects. Table I shows mean age, height, weight, and proportion of abnormal pulmonary function results by gender. Pulmonary function tests were performed in the pulmonary clinic at Ruby Memorial Hospital by trained technicians using a whole body plethysmograph (Model 1085/D, MedGraphics, St. Paul, Minnesota) and spirometer (Model Jaeger MasterScope, VIASYS Healthcare, Hoechberg, Germany) following American Thoracic Society guidelines [16].

D. Cough Measurement Protocol

In order to ensure a repeatable lung volume history prior to a cough, each subject was asked to inhale to total lung capacity, exhale passively to functional residual capacity, inhale a second time to total lung capacity, and hold their breath for 3 seconds. They were then instructed to form a seal with their teeth and lips around the mouth piece connected to the metal tube and were asked to cough. The subject repeated the maneuver 3 separate times with each cough airflow signal being recorded.

E. Feature Extraction

Many features can be derived from a cough airflow waveform. Three example features, including peak airflow, average airflow and peak airflow acceleration, along with the airflow signal versus time are illustrated in Figure 1. A total of 21 features describing the cough airflow pattern were extracted in the time, frequency and joint time-frequency domains. Additionally, the demographic data of gender, age, weight, and height were used as features. Note that two features were used to represent gender. Although it was not tested in this research, it is thought that using two variables rather than one for binary predictors works better for some types of classifier [17], [18]. A complete list of the feature set is presented in Table II. Each feature was normalized to [0 1]. Three cough maneuvers were recorded for

each subject in the study. Repeatability of cough flow features on individuals tested with this recording system was verified in a past study [13].

F. Support Vector Machine Classifier

The classifier implemented in this research was a support vector machine (SVM) with a radial basis function (RBF) kernel. SVM's are known to be robust and provide good generalization performance [19], [20]. The implementation of this SVM depends on two parameters (often referred to as hyperparameters): the SVM soft margin constant, C , and the RBF inverse-width parameter, γ . Both parameters must be tuned to achieve an acceptable trade-off between performance and over-fitting. In this research, we implemented feature selection and hyperparameter determination inside a genetic algorithm (GA) based on the method by Huang and Wang [21] and similar to that proposed by Fröhlich, et al. [22]. The SVM's were programmed using the LibSVM package [23] within Matlab® (2012b, The Mathworks, Natick, MA). All other portions of the software were performed using Matlab. The implementation of the GA is discussed below.

G. Repeated Double Cross Validation

In order to avoid selection bias in the classifier performance evaluation [24], we implemented a repeated ($N=50$) double k-fold ($k=5$) cross validation structure modeled after that of Filzmoser, et al. [3]. Although there are some differences, we have borrowed their nomenclature as well. The software implements the procedure in three nested loops: 1) a repetition loop, 2) an outer cross validation loop, and 3) an inner cross validation loop which is contained within a genetic algorithm. Figure 2 shows an overview of the method.

The outermost loop (the repetition loop) is used to repeat the double k-fold cross validation to assess the variability associated with the particular data segmentation. Within each iteration of the repetition loop, the data are randomly split into five segments on a patient basis. Splitting the data on a patient basis ensures that all three coughs of a single individual remain in the same segment, thereby keeping each segment independent.

For each iteration of the outer cross validation loop, one segment is set aside as the test group. The other four segments are considered the calibration set. The calibration set is then repartitioned (again on a patient basis) into five segments. The calibration set, along with the partition information, is sent to a genetic algorithm. The genetic algorithm (described below) returns values for C and γ as well as the features which have been selected for this iteration. An SVM is trained on the entire calibration set, and tested on the test group. The SVM parameters, the feature set, and the results on the test set (including classification results and associated scores) are all saved for evaluation after completion of all repetitions.

H. Genetic Algorithm

Within the GA, for each generation, the fitness of each genome within the population is assessed by performing a 5-fold cross validation on the calibration set. The fitness scores of each genome were based on the balanced error rate. The folds are fixed, and they are based on the partition definitions passed to the GA from the outer cross validation loop. Each genome was encoded in 58 bits: 16 bits each for C and γ , and 26 bits for the feature mask.

The GA settings included: a population size of 200 with an elite count of 20, a mutation rate of 0.1, and a crossover fraction of 0.7. The GA returns values for C and γ , and selects the feature set which minimizes the prediction error of the entire 5-fold inner cross validation of the calibration set.

I. Assessment of Prediction Accuracy

The prediction accuracy for the classifiers developed in this research was assessed with three measures. The absolute prediction accuracy, P_{abs} , is based on the prediction results from 250 validation sets. Since there were three coughs per subject, majority voting was used to determine the resulting class on a subject basis. The balanced prediction accuracy, P_{bal} was calculated as: $P_{bal} = (P_p + P_n)/2$. Here, P_p is the absolute prediction accuracy from subjects whose class is true positive, and P_n is the absolute prediction accuracy from subjects whose class is true negative. The third measure used in this study was the area under the ROC curve (AUC), which is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance [25].

III. Results

A. Cough Airflow Measurements

Figure 3 shows typical curves representing the airflow and CPEFV relationship during a voluntary cough for a subject with normal spirometry. Figure 4 shows the same curves for a subject with spirometry values that are less than the lower limit of normal for FEV_1 and $FEV_1/FVC\%$. The CPEFV relationship was obtained by plotting airflow generated during a cough versus the volume of air expelled. For the subject with normal spirometry (Figure 3), it can be seen that the flow increased rapidly as the cough approached its peak airflow (11.01 L/sec). The volume of expelled gas reached approximately 0.29 L by peak flow. The total exhaled volume of the cough was approximately 2.02 L. The CPEFV curve shown in Figure 4 was recorded during a cough of a subject with abnormal spirometry. In this case, the maximum flow was lower, 5.03 L/sec, and the total volume of gas expelled during the cough was 0.81 L.

B. Classification

For each of the 50 repetitions, there are five folds in the outer cross validation loop. This results in 250 sets of C , γ , and selected features. Figure 5 shows a histogram of the feature selection results for classification of FEV_1 . Similar results were obtained for FVC and $FEV_1/FVC\%$ classifiers.

The results of the FEV_1 classifier are shown in Figure 6. The FEV_1 prediction accuracy based on majority voting for the entire experiment was 76.65%. The balanced accuracy rate was 76.30%. The area under the ROC curve (AUC) was 79.96%, with lower and upper bounds (95% confidence) of 79.13% and 80.79%, respectively.

The results of the FVC classifier are shown in Figure 7. The FVC prediction accuracy based on majority voting for the entire experiment was 65.36%. The balanced accuracy rate

was 61.00%. The area under the ROC curve (AUC) was 61.27%, with lower and upper bounds (95% confidence) of 60.38% and 62.17%, respectively.

The results of the $FEV_1/FVC\%$ ratio classifier are shown in Figure 8. The $FEV_1/FVC\%$ prediction accuracy based on majority voting for the entire experiment was 76.00%. The balanced accuracy rate was 75.57%. The area under the ROC curve (AUC) was 82.52%, with lower and upper bounds (95% confidence) of 81.53% and 83.5%, respectively.

IV. Discussion

This study analyzed flow patterns from voluntary coughs with a goal of reliably identifying patients with abnormal spirometry. By utilizing a support vector machine classifier a relatively high predictive accuracy was achieved for FEV_1 and $FEV_1/FVC\%$. Identification of subjects with abnormal pulmonary function by analyzing the airflow pattern of their voluntary cough has several advantages. First, it is much less strenuous to cough than to perform conventional spirometry, particularly for young and older subjects. Additionally, tests utilizing cough can be administered much faster and with a minimum of patient training or coaching. These advantages lead to several potential applications for this technique, including the screening of a large population of test subjects in a short period of time with a modest amount of test equipment and a minimum amount of patient training. Another potential use of cough airflow feature analysis is in evaluating the progression or recovery of pulmonary disorders without performing more strenuous testing procedures with more elaborate equipment. For example, with a miniaturized cough measurement system such as the one being developed in our laboratory, this technique could be used for self-monitoring of the onset of asthma, much like current peak flow meters are being utilized.

This research used the equipment described in [13] and builds on the work described in [2]. There are, however, important differences. The current study employed a different type of classifier and used a reduced data set. In [2], features of both cough flow and cough sound were considered. In this article, only cough flow was examined, which is much less technically demanding to measure than cough sound. In addition, the previous work included features measured for an individual subject that were dependent on the entire data set of all subjects. Therefore, the reported prediction accuracy may be too optimistic since the validation would not include completely independent test and training sets. Another difference is that rather than developing a classifier based on patient diagnosis, this study proposed a classifier based on the results of spirometry. Issues related to discrepancies in human interpretation of patient data and history are, therefore, eliminated. Rather, this study is based only on measured data and leaves the interpretation out of the classifier. Subsequently, the classifier results could be used to assist in the interpretation of the respiratory health of the subject.

In this research, care was taken to provide a thorough validation, including assessment of the variability in estimates of the classifier generalization error. The validation process was modeled after that in [3] and included a genetic algorithm for model and feature selection similar to that in [21]. Unlike those techniques however, we incorporated a k-fold cross validation within the GA such that the fitness of every genome in every generation was

evaluated using a complete cross validation. The advantage of this approach is that the SVM hyperparameters are not susceptible to overfitting. The disadvantage is that the method is computationally very demanding. This drawback was mitigated by use of a currently available multiprocessor, multicore computer and the Matlab Parallel Computing Toolbox.

The feature selection over all of the repetitions indicate that there is likely some correlation between many of the features. Figure 5 shows the histogram of features selected for FEV_1 . The histograms for FVC and $FEV_1/FVC\%$ ratio are similar in nature, but vary with respect to most and least often used features. Additionally, none of the features were used in every case, but many were used a similar number of times. We did not investigate the effect of any data reductions techniques (i.e, principal component analysis) on the feature set.

The classification results for FEV_1 and $FEV_1/FVC\%$ were significantly better than the estimates for FVC . This may be due to a lack of information in the cough flow signal or from a lack of volume based features. It appears however, there are a number of things that could be implemented that might improve the results, including an expanded feature set, different SVM kernels, or other classification methods. Finally, this study is limited by a relatively small data set and would be improved by further studies considering healthy populations and those patients with a variety of respiratory conditions.

In summary, this study demonstrates that partial flow-volume relationships generated during voluntary coughs can be used to detect subjects with abnormal spirometry. This measurement is easily performed and does not require expensive equipment. Based on these characteristics, it has the potential for becoming especially useful for mass screening or for testing subjects who are not able to perform conventional pulmonary function testing.

Acknowledgments

The findings and conclusions of this work are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

References

1. Hankinson J, Odencrantz J, Fedan K. Spirometric reference values from a sample of the general u.s. population. *Am J Respir Crit Care Med.* 1999; 159:179–187. [PubMed: 9872837]
2. Abaza A, Day J, Reynolds J, Mahmoud A, Goldsmith W, McKinney W, Petsonk E, Frazer D. Classification of voluntary cough sound and airflow patterns for detecting abnormal pulmonary function. *Cough.* 2009; 5(1):8. [PubMed: 19930559]
3. Filzmoser P, Liebmann B, Varmuza K. Repeated double cross validation. *J Chemometrics.* 2009; 23:160–171.
4. Hammer E, Eber E. *Pediatric Pulmonary Function Testing (Progress in Respiratory Research)*. S Karger AG; 2005.
5. Leith D, Butler J, Sneddon S, Brain J. *Handbook of Physiology. The Respiratory System. Control of Breathing. Section 1, Volume II, part 1.* American Physiological Society; 1986. Cough; 315–336.
6. Barnes P, Gribbin H, Osmanliev D, Pride N. Partial flow-volume curves to measure bronchodilator dose-response curves in normal humans. *J Appl Physiol.* 1981; 50(6):1193–1197. [PubMed: 7263379]
7. Adler S, Wohl M. Flow-volume relationship at low lung volumes in healthy term newborn infants. *Pediatrics.* 1978; 61(4):636–640. [PubMed: 662489]

8. Morgan W, Geller D, Tepper R, Taussig L. Partial expiratory flow-volume curves in infants and young children. *Pediatr Pulmonol.* 1988; 5(4):232–243. [PubMed: 3070472]
9. Wall M, Misley M, Dickerson D. Partial expiratory flow-volume curves in young children. *Am Rev Respir Dis.* 1984; 129(4):557–562. [PubMed: 6711999]
10. Zamel N, Hughes D, Levison H, Fairshter R, Gelb A. Partial and complete maximum expiratory flow-volume curves in asthmatic-patients with spontaneous bronchospasm. *Chest.* 1983; 83(1):35–39. [PubMed: 6848332]
11. Beardsmore C, Wimpres S, Thomson A, Patel H, GP, Simpson H. Maximum voluntary cough: an indication of airway function. *Bull Eur Physiopathol Respir.* 1987; 23(5):465–472. [PubMed: 3450329]
12. Beardsmore C, Park A, Wimpres S, Thomson A, Simpson H. Cough flow-volume relationships in normal and asthmatic-children. *Pediatr Pulmonol.* 1989; 6(4):223–231. [PubMed: 2748217]
13. Goldsmith W, Mahmoud A, Reynolds J, McKinney W, Afshari A, Abaza A, Frazer D. A system for recording high fidelity cough sound and airflow characteristics. *Ann Biomed Eng.* 2010; 38(2): 469–477. [PubMed: 19876736]
14. Thorpe C, Fright W, Toop L, Dawson K. A microcomputer-based interactive cough sound analysis system. *Comput Methods Programs Biomed.* 1991; 36(1):33–43. [PubMed: 1760923]
15. Goldsmith W, Reynolds J, Day J, Kashon M, Frazer D. Prediction of spirometric evaluations based on cough flow. *Am J Respir Crit Care Med.* 2013; (187):A1919.
16. Society AT. Standardization of spirometry: 1994 update. *Am J Respir Crit Care Med.* 1995; 152(3): 1107–36. [PubMed: 7663792]
17. wei Hsu C, chung Chang C, jen Lin C. A practical guide to support vector classification. Tech Rep. 2003
18. Lovell BC, Walder CJ. Support vector machines for business applications. In: Voges K, Pope N, editors *Business applications and computational intelligence.* Hershey, PA, U.S.A: Idea Group; 2006. 267–290.
19. Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery.* 1998; 2:121–167.
20. Ben-Hur A, Weston J. A user's guide to support vector machines. *Methods Mol Biol.* 2010; 609:223–239. [PubMed: 20221922]
21. Huang C, Chieh C. A ga-based feature selection and parameters optimization for support vector machines. *Expert Systems with Applications.* 2006; 31:231–240.
22. Fröhlich H, Chapelle O, Scholkopf B. Feature selection for support vector machines by means of genetic algorithm. *IEEE Comput Soc; Proceedings 15th IEEE International Conference on Tools with Artificial Intelligence;* Dec, 2003. 142–148. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1250182>
23. Chang C-C, Lin C-J. Libsvm: A library for support vector machines. *ACM Trans Intell Syst Technol.* 2011; 2(3) aticle No. 27. [Online]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
24. Cawley G, Talbot N. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J Mach Learn Res.* 2010; 11:2079–2107.
25. Fawcett T. An introduction to ROC analysis. *Pattern Recogn Lett.* Jun; 2006 27(8):861–874. [Online]. Available: DOI: 10.1016/j.patrec.2005.10.010

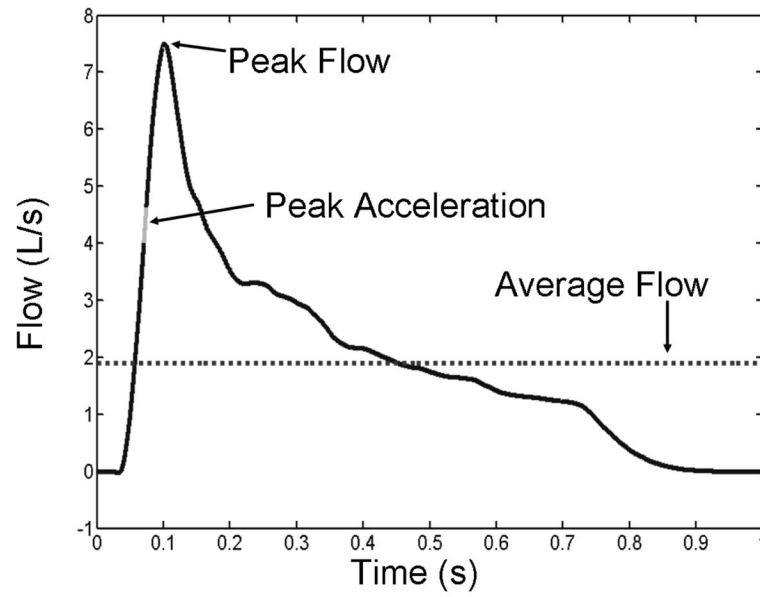


Fig. 1. Airflow during a cough as a function of time for a subject with normal spirometry. Three example features are illustrated: peak flow, average flow, and peak acceleration.

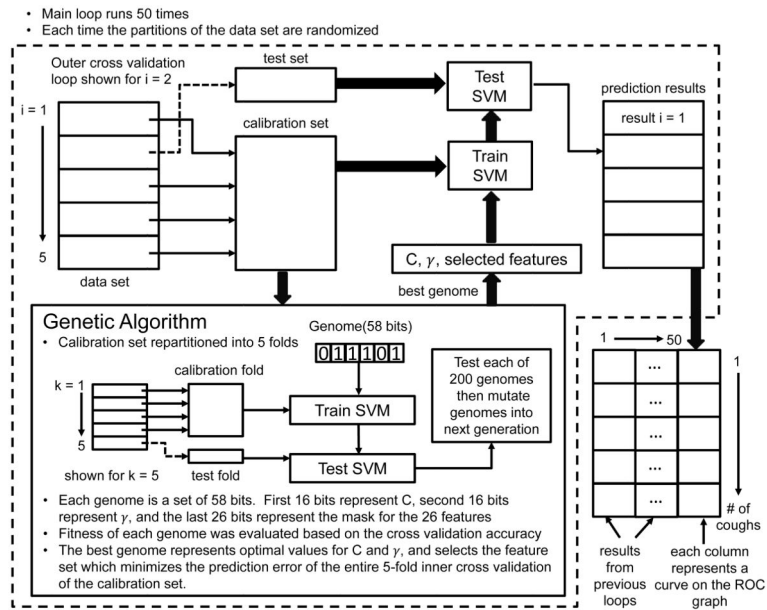


Fig. 2.
Classifier diagram.

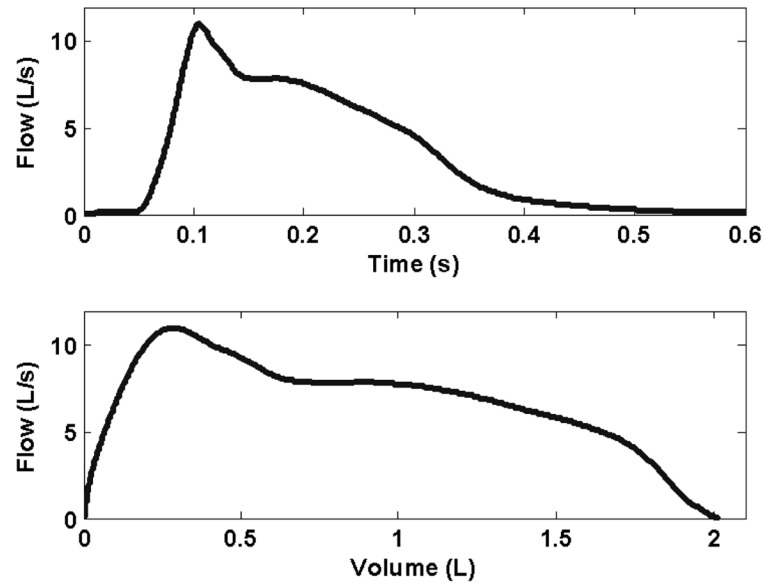


Fig. 3.
An example airflow waveform (upper panel), and a partial expiratory flow-volume curve (lower panel) recorded during a voluntary cough for a subject with normal spirometry.

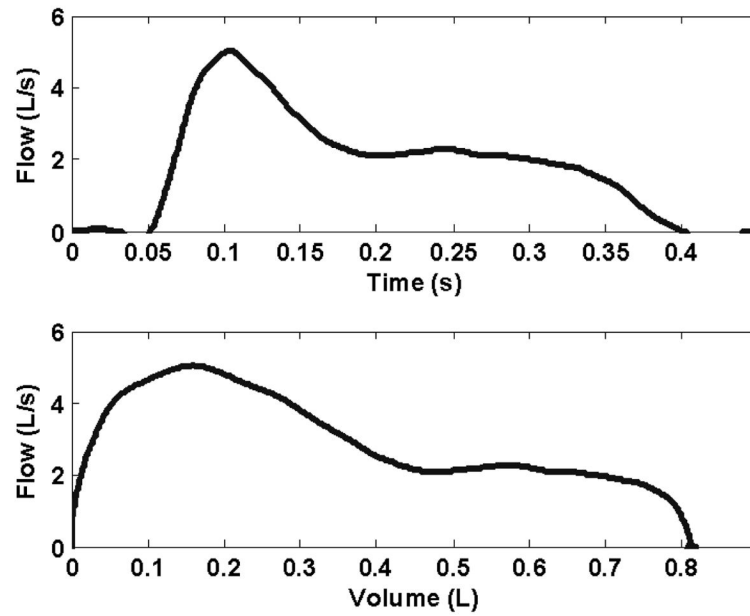


Fig. 4.

An example airflow waveform (upper panel), and a partial expiratory flow-volume curve (lower panel) recorded during a voluntary cough for a subject whose spirometry was less than the lower limit of normal for FEV_1 and $FEV_1/FVC\%$.

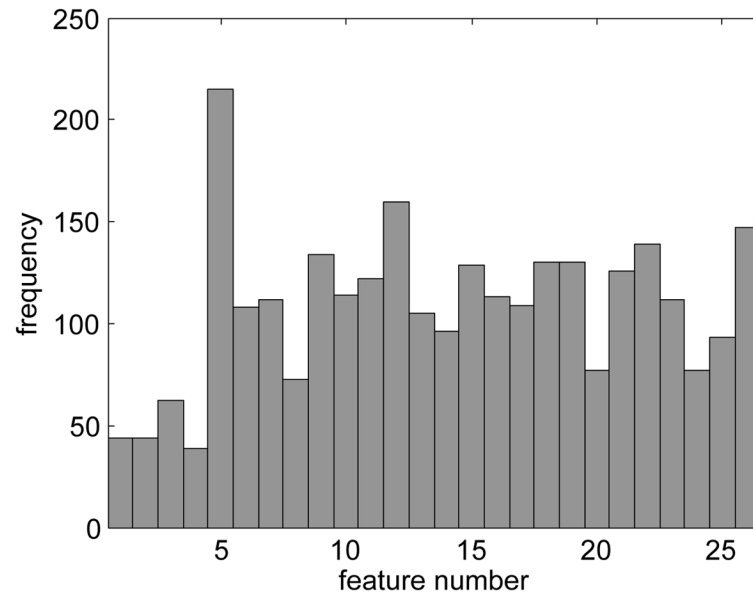


Fig. 5. Histogram showing the number of times a particular feature was selected for the FEV_1 classifier.

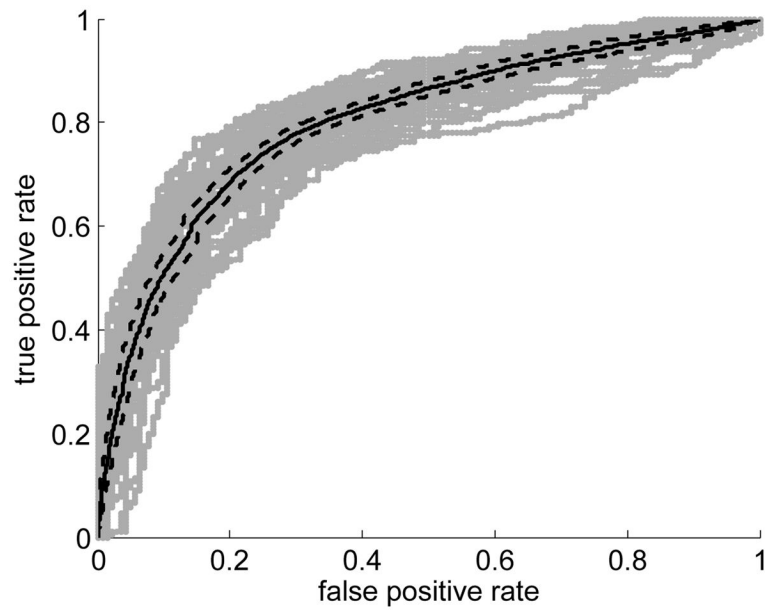


Fig. 6. ROC curve (solid black line) for the FEV_1 classifier. The grey lines represent the ROC curves calculated on each of the individual repetitions. The dashed black lines represent the envelope of the bounding boxes created from the pointwise confidence bounds (95% confidence).

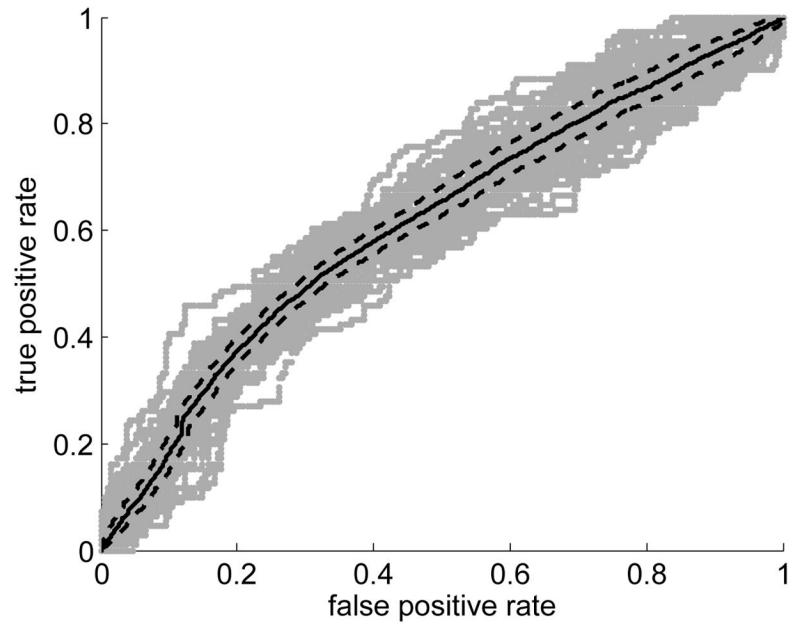


Fig. 7. ROC curve (solid black line) for the *FVC* classifier. The grey lines represent the ROC curves calculated on each of the individual repetitions. The dashed black lines represent the envelope of the bounding boxes created from the pointwise confidence bounds (95% confidence).

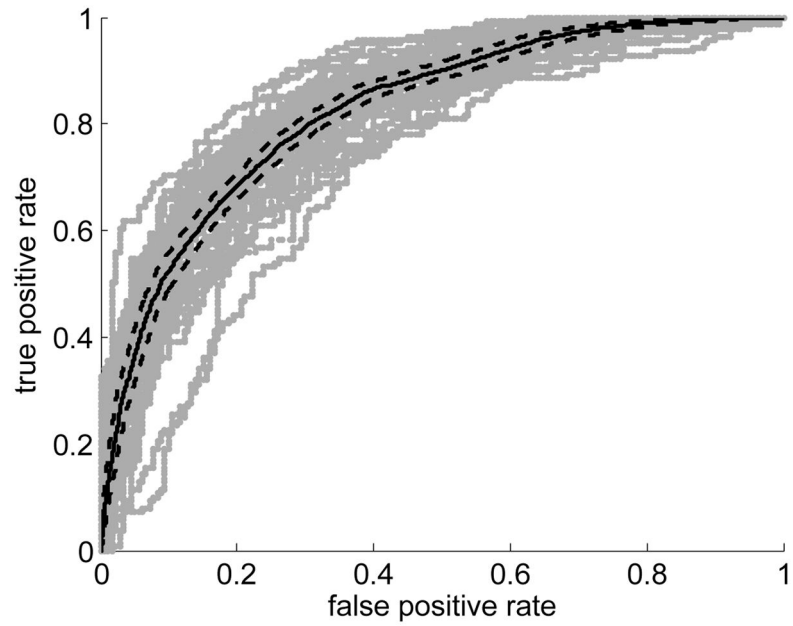


Fig. 8. ROC curve (solid black line) for the $FEV_1/FVC\%$ classifier. The grey lines represent the ROC curves calculated on each of the individual repetitions. The dashed black lines represent the envelope of the bounding boxes created from the pointwise confidence bounds (95% confidence).

TABLE I

Description of group populations of test subjects

	Male (n = 56)	Female (n = 51)
age (mean \pm SD, years)	54.96 \pm 14.15	54.78 \pm 15.58
height (mean \pm SD, cm)	175.38 \pm 8.86	160.45 \pm 6.58
weight (mean \pm SD, kg)	91.20 \pm 26.32	80.48 \pm 25.75
<i>FEV</i> ₁ LLN	24	24
<i>FEV</i> ₁ < LLN	32	27
FVC LLN	18	19
FVC < LLN	38	32
<i>FEV</i> ₁ / <i>FVC</i> % LLN	26	34
<i>FEV</i> ₁ / <i>FVC</i> % < LLN	30	17

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE II

Demographic Features and Features Extracted from Cough Airflow Signals

Feature #	Description
1	male: this is 1 for male gender, 0 otherwise
2	female: this is 1 for female gender, 0 otherwise
3	age (years)
4	weight (kg)
5	height (cm)
6	peak cough flow (L/s)
7	average cough flow (L/s)
8	maximum cough flow acceleration (L/s^2)
9	total cough volume (L)
10	time at which 25% cough volume has been expelled/time at which 100% cough volume has been expelled
11	time at which 50% cough volume has been expelled/time at which 100% cough volume has been expelled
12	time at which 75% cough volume has been expelled/time at which 100% cough volume has been expelled
13	25% total time of cough/cough volume
14	50% total time of cough/cough volume
15	75% total time of cough/cough volume
16	time at peak cough flow/total time of cough
17	crest factor: peak cough flow/Root Mean Square (RMS) of cough flow
18	form factor: RMS of cough flow/mean cough flow
19	transit time: $\int \frac{\text{cough flow}}{\text{total volume}} \cdot t \cdot dt$ (s)
20	skewness: $\frac{E(x - \mu)^3}{\sigma^3}$ where μ , and σ are the mean, and the standard deviation of the cough flow signal respectively
21	kurtosis: $\frac{E(x - \mu)^4}{\sigma^4}$ where μ , and σ are the mean, and the standard deviation of the cough flow signal respectively
22	beta: the inverse power law $1/f^\beta$ of the power spectrum
23	cough flow variance
24	cough flow variance normalized with respect to volume
25	wavelet parameter based on the variability in the wavelet detail coefficients found in the wavelet decomposition of the cough flow
26	cough length (s)