ISSN No. 0976-5697

Volume 3, No. 1, Jan-Feb 2012

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

Spirometry Modelling: A New Concept of Post Processing of Traditional Data

Ms. Prachi D. Junwale *
M. Tech IV Sem, Department of Computer Science & Engg,
Government College of Engineering, Amravati
Amravati (M.S.) - 444604, India.
prachi.junwale@gmail.com

Prof. A. W. Bhade
Head, Department of Information Technology,
Government College of Engineering, Amravati
Amravati (M.S.) - 444604, India.
bhade.archana@gcoea.ac.in

Dr. P. N. Chatur
Head, Department of Computer Science & Engg,
Government College of Engineering, Amravati
Amravati (M.S.) - 444604, India.
chatur.prashant@gcoea.ac.in

Abstract: Respiratory disease is a medical term that encompasses pathological conditions affecting the organs and tissues. A lungs disease affects health condition of many people. Lungs diseases can be curable in early detection. Spirometry is valuable for diagnosing specific lung disorders as well as detecting lung disease at an early stage. Spirometry (the measuring of breath) is the most common of the pulmonary function tests and uses an instrument called a spirometer to measure the amount of air entering and leaving the lungs. This test is often used to help doctors diagnose and determine the severity of various respiratory diseases. The output of spirometry is in the form of graphs i.e. flow-volume loop and volume-time curve. This graph gives various parameters that are used for spirometry modelling. Spirometry modeling gives flow–time curves. This curve is modeled using statistical data mining technique, which is used to obtain new information about breathing condition which makes distinguish between healthy and diseased subjects.

Keywords: respiratory system; lung disorders; spirometry; modelling; statistical data mining.

I. INTRODUCTION

The respiratory system plays a vital role in delivering oxygen to the body i.e. fuel for all the body's functions. It also removes carbon dioxide waste, eliminates toxic waste, regulates temperature, and stabilizes blood acid-alkaline balance (pH). The lungs are the largest part of the respiratory system and have both "respiratory" and "non-respiratory" functions. The respiratory function involves gas exchange i.e. the transfer of oxygen from the air into the blood and the removal of carbon dioxide from the blood [2]. Non-respiratory lung functions are mechanical, biochemical, and physiological. The lungs provide a defense against bacterial, viral and other infectious agents, remove various metabolic waste products, control the flow of water, ions, and large proteins across its cellular structures, and manufacture a variety of essential hormones and chemical agents that have important biological roles.

Respiratory diseases can arise from a number of causes, including inhalation of toxic agents, accidents, and harmful lifestyles, such as smoking. Infections, genetic factors, and anything else that affects lung development, either directly or indirectly, can cause respiratory symptoms. Lung disease is any disease or disorder that occurs in the lungs or that causes the lungs to not work properly. Various techniques are used to detect respiratory problem such as pulmonary function tests (PFT). PFT measure how efficiently lungs perform. PFTs help physicians to evaluate lung function. Spirometry is the most widely used pulmonary function test.

A. Spirometry:

Spirometry is the most common of the lung function tests. These tests look at how well your lungs work. Spirometry shows how well you breathe in and out. Breathing in and out can be affected by lung diseases such as chronic obstructive pulmonary disease (COPD), asthma, pulmonary fibrosis and cystic fibrosis [5]. Spirometry is valuable for diagnosing specific lung disorders as well as detecting lung disease at an early stage [5]. In summary, spirometry can:

a. Identify the presence of lung disease.
b. Help the physician assess the severity of the patient's lung disease.
c. Identify pulmonary disease symptoms and the degree of disability.
d. Assist in the management of patients with lung disease.
e. Provide early detection of pulmonary disease.
f. Assist in convincing patients to quit smoking.
g. Help the physician assess the effects of therapy or medications.

Spirometry measures the amount (volume) and/or speed (flow) of air that can be inhaled and exhaled. The most common parameters of spirometry test are:

a) Forced expiratory volume in one second (FEV1): This is the amount of air you can blow out within one second. With

© 2010, IJARCS All Rights Reserved

normal lungs and airways you can normally blow out most of the air from your lungs within one second.

b) Forced vital capacity (FVC): The total amount of air that you blow out in one breath.

c) FEV1 divided by FVC (FEV1/FVC): This ratio shows the total amount of air that you can blow out in one breath.

Two curves are shown after the spirometry tests that are flow-volume loop and volume-time curve [5]. The various parameters such as PEF (peak expiratory flow), FIT (forced inspiratory time) and FET (forced expiratory time) can be determined by using flow-volume loop and volume-time curve. These parameters are used to determine the values of model parameters. PEF is the maximal expiratory flow rate achieved and this occurs very early in the forced expiratory manoeuvre. FET is the forced expiratory time. FIT is the forced inspiratory time.

In this paper, we will concentrate on data mining approach. For spirometry modelling, statisitca data miner is used to model flow–time curves.

B. Statistical Data Mining:

Data mining has become a common activity in practically all areas of business, research and government. The amount of data that is available in various data bases, data warehouses, or via automated data collection equipment is quickly growing larger and becoming ever more complex [4]. Computer hardware (architectures), memory, and processor speeds make it possible to apply novel advanced algorithms to extract information about complex interactions, clusters, and other systematic patterns in the available information. STATISTICA Data Miner was built from advanced modern software tools, and incorporates all standard as well as the most cutting-edge methods and algorithms for extracting information from data [4]. The program also offers very flexible graphical tools that in practice will prove indispensable for data exploration and visual data mining and for interpreting the "nuggets of information" extracted from the data.

STATISTICA offers facilities to convert almost any data set into a wide variety of graphical representations via a large selection of graph definition and customization options. From instantly produced, predefined graphs to enhance the interpretation of specific statistical analyses, to graphs designed for rapid data exploration and graphs for detailed custom presentation of data, STATISTICA has a graphing method available to meet every type of analysis. It has a critical importance with regard to the goals of data mining project [4].

II. RELATED WORK

There are various fields that are applied on spirometric data for performing post-processing research in respiratory disease analysis area. Such field includes pattern recognition approach [2], image processing, genetic algorithm [10], and artificial neural network [9].

One of such a research on spirometry data is "automatic classification of spirometric data" that was implemented in 1979. In this, pattern recognition principles have been applied to 200 sets of spirometric data obtained from pulmonary function laboratory patients. Each patient was classified by a pulmonary specialist as normal, restricted, or mildly, moderately, severely, or very severely obstructed [2]. Each patient was represented by a five-element pattern vector consisting of forced vital capacity (FVC), forced expiratory volume in one second (FEVI), mid-maximum flow rate (MMFR), and flow rates with 50 and 25 percent of the vital capacity remaining (PSO and P25) normalized by predicted values [2]. By Karhunen-Loeve expansion techniques, this vector was reduced to a two-feature pattern vector with only a 6 percent residual mean square representation error. The more important feature essentially represented the average of the three flow rates, while the second feature depended on FVC and FEV1. Data were divided into training and testing sets, and using the former, a parametric Bayes classifier and one- and two-layer pairwise Fisher linear classifiers, were designed to assign patterns described by the two derived features to one of the six categories. With the testing set, overall recognition rates were 81 to 82 percent, with most errors representing misclassifications within the four obstructive categories. If the four obstructive classes were considered as a single class, the recognition rate increased to about 94 percent [2].

Spirometric data is also used for Case Finding, Diagnosis, and Management of Chronic Obstructive Pulmonary Disease (COPD) [6]. Chronic Obstructive Pulmonary Disease (COPD) is a leading cause of morbidity and mortality. COPD is diagnosed in symptomatic individuals through spirometric testing demonstrating irreversible airflow obstruction [5]. Spirometry in primary care settings for case finding, diagnosis, and management in all adults with persistent respiratory symptoms or having a history of exposure to pulmonary risk factors is controversial [6]. The main objective is to conduct a systematic review to determine:

a. the prevalence of COPD and airflow obstruction;
b. if spirometry improves smoking cessation;
c. if effectiveness of COPD therapies varies based on baseline or change in spirometric severity;
d. whether spirometry provides independent prognostic value related to pulmonary outcomes.

Spirometry, in addition to clinical examination, improves COPD diagnostic accuracy compared to clinical examination alone. The primary benefit of spirometry is to identify individuals who might benefit from pharmacologic treatment in order to improve exacerbations. These include adults with symptomatic, severe to very severe airflow obstruction. Widespread spirometric testing is likely to label a large number of individuals with disease and result in considerable testing and treatment costs and health-care resource utilization.

The research area of post-processing of spirometric is done by using the genetic algorithm principle [10]. Spirometry deals with finding and predicting respiratory system pathologies through instrumentation that mainly carries out measurements on the volume and the air flow expired from lungs. In many cases, during spirometric and pneumotachographic trials in hospital, there are people who are not able to begin or to complete their tests because of diverse difficulties due to presumable pathologies [10]. Hence, these trials may be lost if they are not recovered and postprocessed in adequately, at

© 2010, IJARCS All Rights Reserved

least to display the expiration trend and step. This module gives rapid techniques of helping physiopathologists to extract information from a non-complete expiration curve as spirometric postprocessing. The two techniques are based on work of breath (WOB) and controlled genetic algorithm (CGA), respectively [10]. A comparison is performed between the two techniques; the WOB is calculated by assuming classes of fixed resistance R according to the age, to the sex, to the previous pathologies, etc., while the CGA technique provides a strict monitoring of GA steps in order to reduce uncertainty of final results.

In this paper, we will concentrate on statistical data mining approach for spirometry modeling.

III. SPIROMETRY MODEL

Spirometry lung functioning test is based on the flow measurement. During spirometry test, the maximal expiration and the maximal inspiration are essential things. The flow–time curves can be modeled on the basis of the measurements [8]. Fig. 1 shows spirometry model that is used to describe the airflow during the maximal inspiration *Qin*(*t*) and during the maximal forced expiration *Qout*(*t*). The spline functions were used for the modeling process [7]. The spline functions consist of piecewise lengths of regression functions that give the best fit to localized sections of the data. Statistica data miner concept is used for regression function [4]. The flow -time regression functions were fitted to the flow–time curve based on the spirometry measurements. The maximal inflow *Qmaxin (t)* has been modeled with the regression function given as in equation (1).

$$Qmaxin(t) = A_{in} . \sin(\bar{\omega}. t), \quad t0 \le t \le t1. \tag{1}$$

Where A *in* is amplitude, ϖ is pulsation, (t0, t1) the time of breath–in [3]. The maximal outflow *Qmaxout (t)* has been modeled by two exponential spline functions as shown in equation (2).

$$Qmaxout(t) = \begin{cases} A^1_{out}(1 - e^{-B1out\, t}), & t1 < t \le t2. \\ A^1_{out} . e^{-B2out\, t}, & t2 < t \le t3. \end{cases} \tag{2}$$

Where *A*1*out*, *A2out*, *B*1*out*, *B2out* are the parameters of regression function [1]. Parameters of the model, which is described by above two equations, are arranged in the vector p [p1,p2,…,pi]= [*Ain*, ϖ, *A1out*, *B1out*, *A2out*, *B2out*], where *i* = 1,2...6 and the spirometry measurement results are arranged in the vector of measurement y = [*y*1 , *y*2 ,..., *y j*]= [*FVC*, *PEF*, *PIF*,*TPEF*, *FET*], where *j* = 1,2...5 .

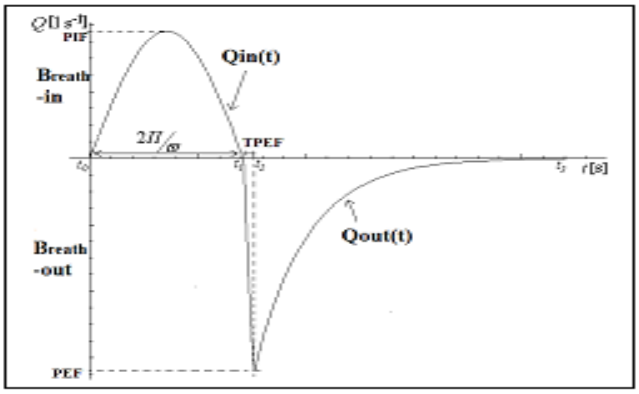

Figure 1. The flow–time curves: *Qin* (*t*) and *Qout* (*t*)

The parameters of model were calculated on the basis of traditional spirometric parameters. The flow-time regression functions, given by the equations (2) and (3), were fitted to the flow-time curve *Q*(*t*) obtained as the spirometry measurement result. The fitting procedure, by adjusting the parameters of regression functions, assures the equality of the parameters *FEV*1, *FVC*, *PEF*, *FIT*, *FET* calculated on the basis of the spirometry measurements with the same ones resulting from the model functions. This implies that the traditional spirometric parameters such as *FEV*1, *FVC*, *PEF*, *FIT* and *FET* are used to determined ϖ, *A1out*, *B1out*, *A2out*, *B2out* model parameters respectively with 1% difference. The model parameter Ain is constant and is equal to 3.912 for normal and disease cases. We assume that the parameter *FET* = 6 *sec.* (which is the minimal, correct value) and the ratio of *FIT: FET* = 1:2.

Spirometry model technique is used to obtain new information about breathing condition which makes distinguish between healthy and diseased subjects.

IV. CONCLUSION

The various spirometric parameters such as FVC, FEV1, PEF, FIT and FET are used to determined spirometric model parameters. The spirometric model is designed on the basis of model parameters using spline functions. Statistical data mining concepts are used for modeling. The various model parameters such as ϖ, *Aout*1 and *Aout*2, when compared to the traditional diagnostic parameters, give new information concerning breathing conditions.

V. REFERENCES

[1] P. Quanjer, J. Tammeling, J. Cotes, "Long volumes and forced ventilatory flows: report of working party, standardization of lung function tests", Eur. Respir. Journal, vol. 6, 1993, pp. 5-40.

[2] P. Quanjer, M. Lebowitz, I. Gregg, "Peak expiratory flow: conclusions and recommendations of a Working Party of the European Respiratory Society. Official ERS Statement", Eur. Respir. J., vol. 10, 1997, pp. 2-8.

[3] J. Kowalski, A. Koziorowski, L. Radwan, Ocena czynności płuc w, chorobach układu oddechowego, "The estimation of the lung function at the pulmonary diseases", Warszawa, Borgis, 2004.

[4] e-handbook of Statistica **http://www.statsoft.com** 2005

© 2010, IJARCS All Rights Reserved

[5] e-handbook of breeze-suite http://www.medgraphics.com 2005

[6] Timothy J. Wilt, Dennis Niewoehner "Use of Spirometry for Case Finding, Diagnosis, and Management of Chronic Obstructive Pulmonary Disease (COPD)", AHRQ Publication No. 05-E017-2, 2005.

[7] Renata Kalicka, Wojciech Słomiński and Krzysztof Kuziemski, "Modelling of Spirometry: diagnostic usefulness of model parameters", IEEE International conference on Computer as a Tool, 2007.

[8] R. Kalicka, W. Słomiński, K. Kuziemski "Modelling of the spirometry measurements", XV Krajowa Konferencja Naukowa Biocybernetyka i Inżynieria Biomedyczna, Wrocław, 2007, pp. 100-104.

[9] S. Jafari, H. Arabalibeik and K. Agin, "Classification of normal and abnormal respiration patterns using flow volume curve and neural network", IEEE international conference on Information Technology, 2009.

[10] Aimé Lay-Ekuakille, Giuseppe Vendramin, and Amerigo Trotta, "Spirometric Measurement Postprocessing: Expiration Data Recovery", presented at *IEEE sensor journal*, vol.10, no. 1, 2010.

© 2010, IJARCS All Rights Reserved