

Speech Signal Analysis as an Alternative to Spirometry in Asthma Diagnosis

Kutor John*, Srinivasan Balapangu, Adofo K Jeromy, Dellor Atsu Albert, Nyakpo Christopher and Brown Akwetey Godfred

Department of Biomedical Engineering, University of Ghana, Ghana

Abstract

Speech production involves the vibration of the vocal cords. Voice changes will however occur in asthma due to the inflamed lung airways. Spirometry is a well-known technique employed in diagnosis of asthma to give information on patient pulmonary function. The purpose of this research was to investigate the correlation between FEV_1/FVC (Forced Expiratory Volume to Forced Vital Capacity) ratio obtained from spirometry and Harmonics-to-Noise Ratio (HNR) obtained from human speech, in order to determine whether speech analysis could be an alternative to spirometry in diagnosing asthma. Spirometry data was obtained from 150 subjects, who were asthmatic patients attending the Korle-Bu Teaching Hospital, Ghana. Speech data consisting of the vowel sounds /a:/, /e:/, /ɛ:/, /i:/, /o:/, /ɔ:/, /u:/, consonant /s:/ and phrase "She sells", was also recorded from the subjects. 33 samples were selected and analyzed to generate speech parameters with Praat software. Correlation was established between HNR from the speech signals and spirometry data FEV_1/FVC . The highest correlation coefficient was observed between HNR and vowel sound /ɛ:/ (42.08%). In conclusion, among the other speech vowels and phonemes, Harmonics-to-Noise ratio (HNR) of /ɛ:/ sound showed the most promise to being a suitable alternative to spirometry in asthma diagnosis.

Keywords: Harmonics-to-noise ratio; FEV_1 ; FVC; Asthma; Speech; Diagnosis

Introduction

Sound production in humans involves airflow from the lungs through the larynx, the vibration of the vocal cords and resonance in the oral and nasal cavities [1]. Infections that cause physical changes in any of the parts of the pathway of sound production therefore have the tendency to affect the natural sounds an individual produces during speech [2]. There has been research into human voice analysis in diagnosis medical conditions that affects voice parameters such as depression, schizophrenia and autism spectrum disorders [3]. This paper investigates the effect of asthma on an individual's speech and the possibility of diagnosing asthma via speech.

Asthma is a chronic immune inflammatory disorder influenced by many factors. In 2007, it was stipulated that about 300 million people suffer from asthma [4]. Another report in 2014 also stated that about 334 million people from all ages suffer from asthma, with the most prevalence of symptoms among 18-45 years old [5]. Asthma is an obstructive lung disease. A recognizable effect of bronchial asthma is the inflammation of the expiratory organs such as the trachea, bronchi, bronchioles and alveoli with wheezing being a key symptom during asthmatic episodes [5]. Other symptoms of asthma include coughing, shortness of breath and chest tightness. The airways of asthmatic patient are hypersensitive to stimulus and allergies causing a chronic inflammation of the airways when exposed to such triggers.

Forced spirometry is a pulmonary function test that is used in medical evaluation of patients complaining of shortness of breath. It is often used in assessment and diagnosis of asthma. It measures the efficacy of airflow into and out of the lungs (inhalation and exhalation). Forced vital capacity (FVC) and Forced Expiratory Volume in one second (FEV_1) are key parameters that are obtained from spirometry [6,7]. The ratio FEV_1/FVC , also known as the *Tiffeneau-Penelli* index, is used in the diagnosis of obstructive and restrictive lung diseases [8-10]. Global Initiative for Chronic Obstructive Lung Disease (GOLD), recommends using a post-bronchodilator FEV_1/FVC ratio of less than 0.7 to define an irreversible air-flow limitation and thus present an

indication of the presence of disease [10]. The FEV_1/FVC ratio is therefore a suitable reference against which speech parameters may be correlated to obtain a relationship between asthmatic condition and speech.

Voice analysis involves the extraction of parameters such as harmonics-to-noise ratio (HNR), jitter, shimmer, formant frequency etc. from voice signals to determine its characteristics, which may be used for applications including speech recognition and disease diagnosis [11,12]. Other research established that there is a difference between voice parameters of asthmatic patients and non-asthmatic patients [13,14].

Harmonics-to-noise ratio (HNR) describes the degree of acoustic periodicity in a signal, that is, how much of the energy of the signal is in the periodic part of the signal as compared to the noise in the signal. In a study, HNR was found to be a good index for degree of hoarseness [15]. For this study, FEV_1/FVC ratio was correlated with HNR because an initial analysis of the data showed that HNR was a more sensitive index of vocal function than the other speech parameters. This research therefore seeks to investigate the correlation between FEV_1/FVC and the speech parameters HNR, in an attempt to find out whether speech is a possible alternative to spirometry in the diagnosis of asthma.

Methodology

Subjects

Spirometry and speech data of 150 asthmatic patients at the Korle-Bu Teaching Hospital (Ghana) were taken, along with other information such as age, mass, height and sex. The age range was 18 to

*Corresponding author: Kutor John, Department of Biomedical Engineering, University of Ghana, Ghana, Tel: +233243147354; E-mail: jkutor@ug.edu.gh

Received September 12, 2018; Accepted September 25, 2018; Published September 30, 2018

Citation: John K, Balapangu S, Jeromy AK, Albert DA, Christopher N, et al. (2018) Speech Signal Analysis as an Alternative to Spirometry in Asthma Diagnosis. J Biomed Eng Med Devic 3: 136. doi: 10.4172/2475-7586.1000136

Copyright: © 2018 John K, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

45 years. Consent was obtained from subjects before including them in the research. Data from 33 of these patients was then taken for analysis, since the other data had speech errors or had incomplete information.

Acoustic data extraction

The speech signals to be analyzed were continuous and sustained pronunciation of vowels sounds: /a:/, /e:/, /ɛ:/, /i:/, /o:/, /ɔ:/, /u:/, consonant /s:/ and phrase ‘She sells’. The vowel sound notation used is based on the IPA (International Phonetic Alphabet) symbols. Patients were made to pronounce each of these sounds and the phrase three consecutive times while being recorded at a sampling frequency of 44.1 kHz. The recorder used was Sony ICD px333 Voice Recorder. Speech parameters were then extracted using Praat acoustic analysis software version 6.0.08. The speech data were filtered at a gain of 40 dB (where

necessary) using Adobe Audition CS6 in order to attenuate background noise. Figure 1 shows a sample audio file opened in the Praat interface and Figure 2 shows the view of a selected vowel sound segment from the full audio file shown in Figure 1.

Statistical analysis

The extracted speech parameters and spirometry data were exported to Microsoft Excel 2013, where regression analysis was carried out to establish a correlation between HNR of the various sounds and the spirometry parameter FEV₁/FVC ratio. A scatter plot was done, and linear and polynomial regression analysis between HNR and FEV₁/FVC ratio was performed for the vowels sounds /a:/, /e:/, /ɛ:/, /i:/, /o:/, /ɔ:/, /u:/, and the phrase “she sells” and the coefficient of determination (R²) values were noted (Figure 3).

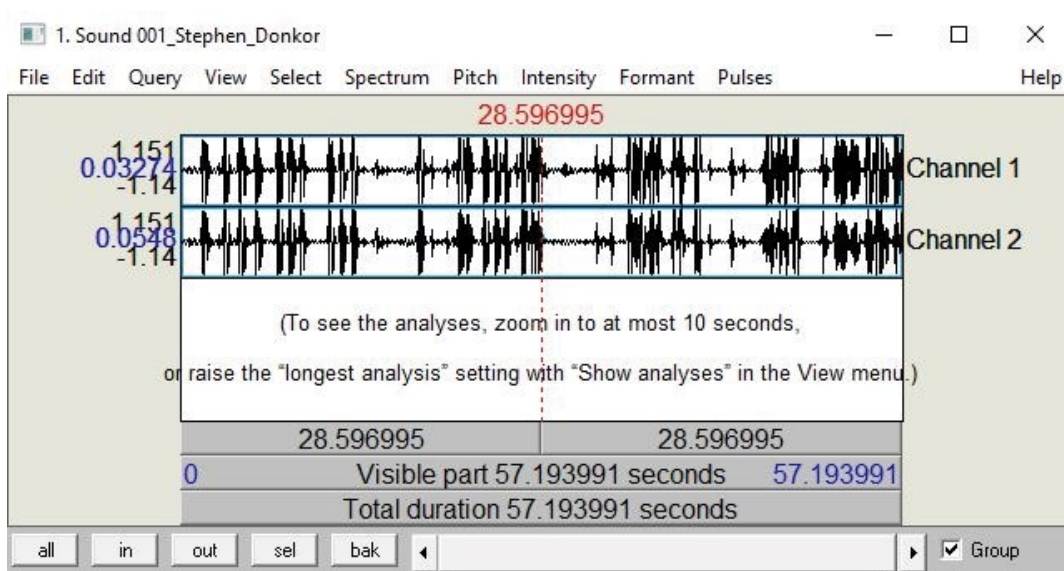


Figure 1: Praat user interface showing an opened audio file.

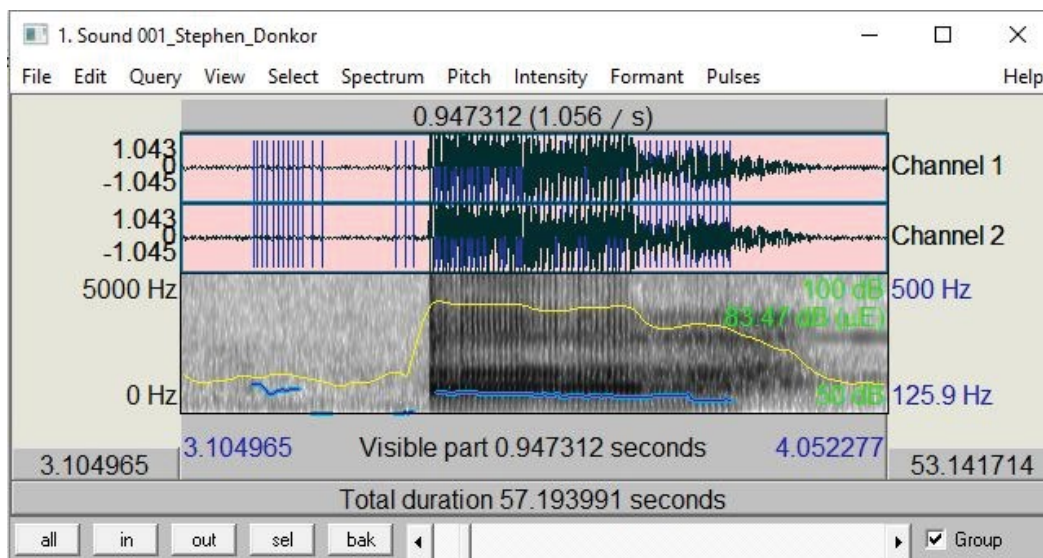


Figure 2: Selected vowel section in Praat interface.

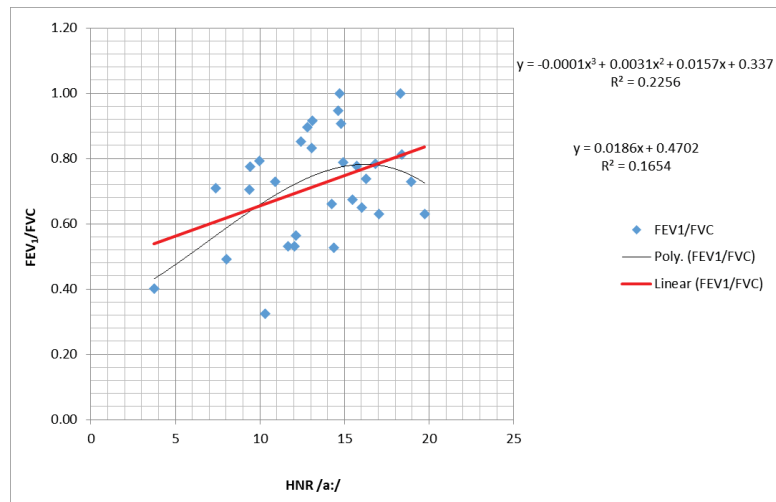


Figure 3: Plot of FEV₁/FVC vs HNR obtained from /a:/ sound.

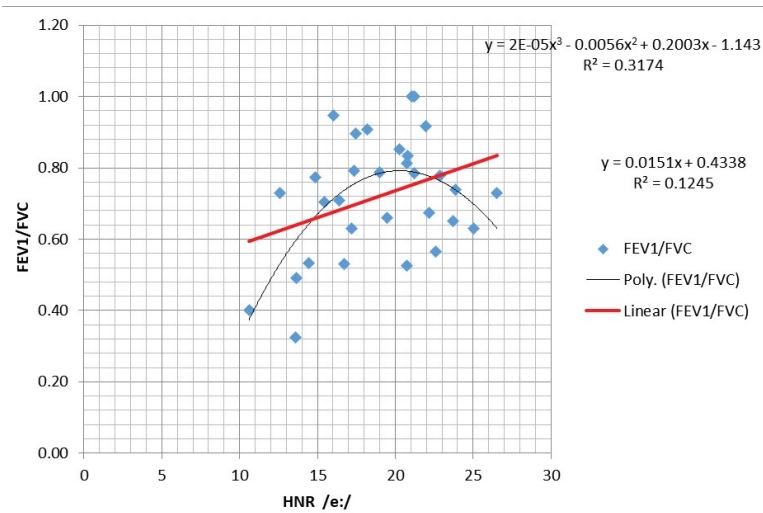


Figure 4: Plot of FEV₁/FVC vs HNR obtained from /e:/ sound.

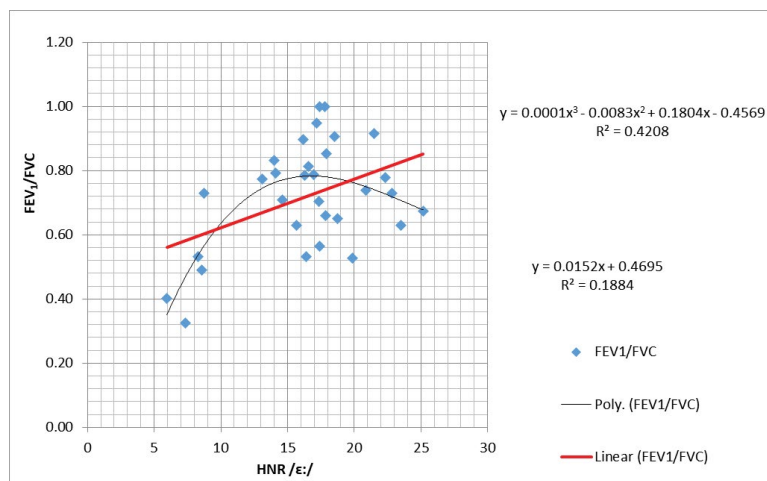


Figure 5: Plot of FEV₁/FVC vs HNR obtained from /ɛ:/ sound.

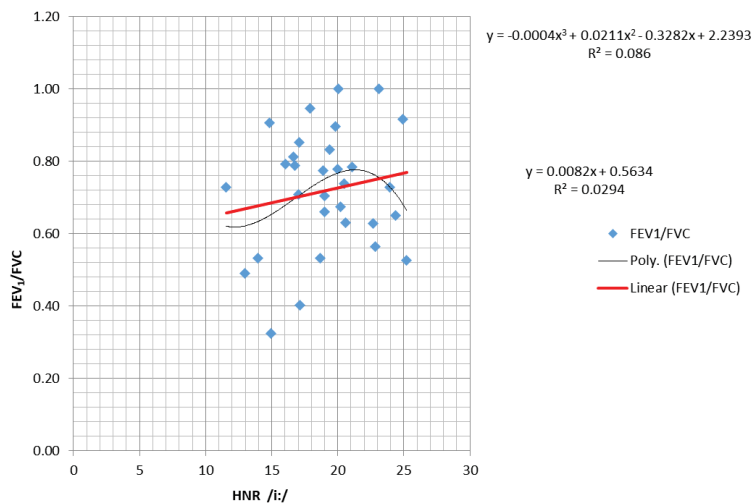


Figure 6: Plot of FEV₁/FVC vs HNR obtained from /i:/ sound.

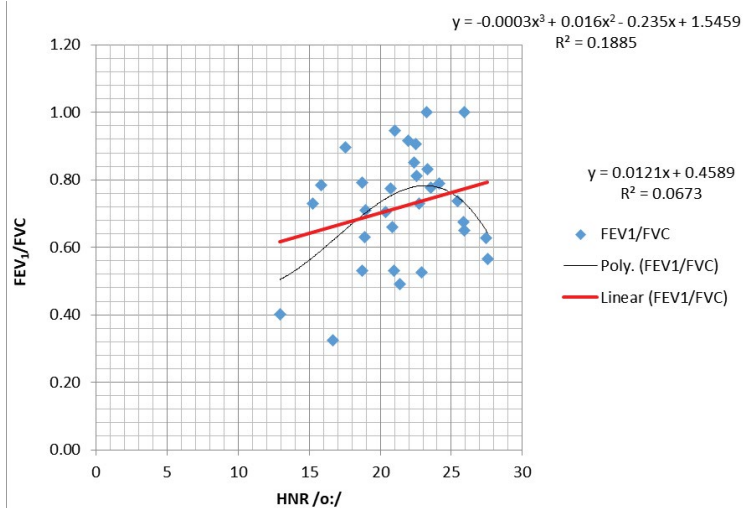


Figure 7: Plot of FEV₁/FVC vs HNR obtained from /o:/ sound.

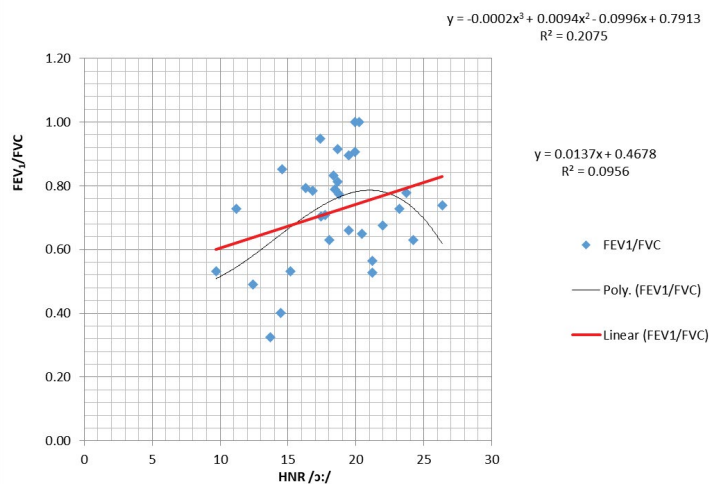


Figure 8: Plot of FEV₁/FVC vs HNR obtained from /ɔ:/ sound.

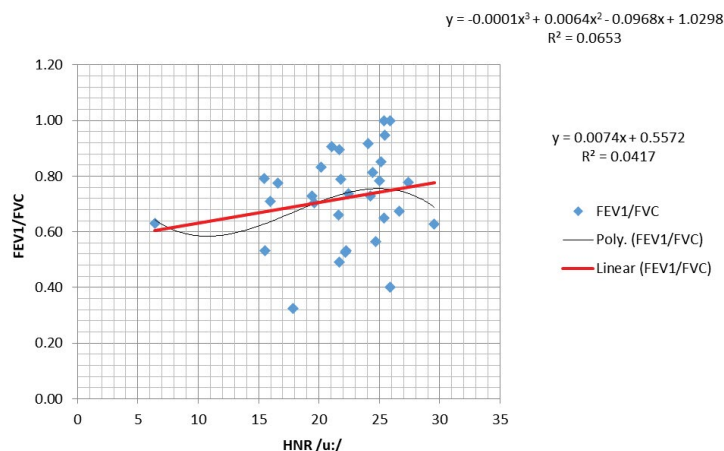


Figure 9: Plot of FEV_1/FVC vs HNR obtained from /u:/ sound.

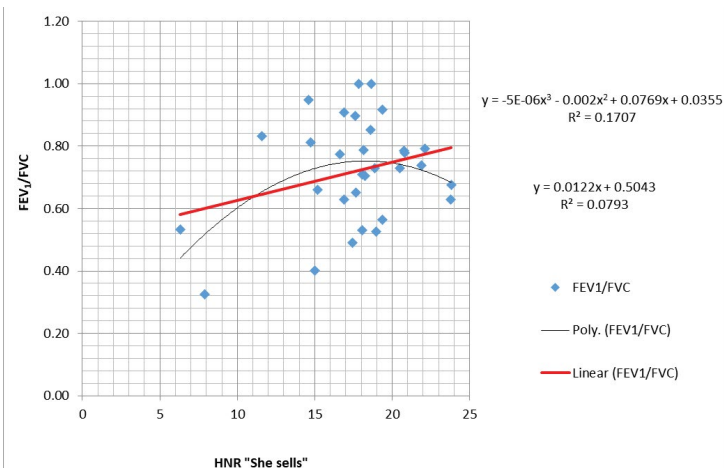


Figure 10: Plot of FEV_1/FVC vs HNR obtained from "She sells" phrase.

Results and Discussion

The vowel sound /a:/ presented an R^2 value of 22.56% in correlating HNR with FEV_1/FVC ratio with a cubic polynomial regression. The equation obtained was:

$$y = -0.0001x^3 + 0.0031x^2 + 0.0157x + 0.337 \quad (3)$$

The linear regression equation obtained was (Figure 4):

$$y = 0.0186x + 0.4702 \quad (4)$$

and it yielded an R^2 of 16.54%, where $y = FEV_1/FVC$ and $x = HNR /a:/$

The vowel sound /e:/ presented an R^2 value of 31.74% in correlating HNR with FEV_1/FVC ratio with a cubic polynomial regression. The equation obtained was:

$$y = 2 \times 10^{-5}x^3 - 0.0056x^2 + 0.2003x - 1.143 \quad (15)$$

The linear regression equation obtained was (Figure 5):

$$y = 0.0151x + 0.4338 \quad (16)$$

and it yielded an R^2 of 12.94%,

where $y = FEV_1/FVC$ and $x = HNR /e:/$

The vowel sound /ε:/ presented the highest R^2 value between HNR and FEV_1/FVC ratio compared to the results from the other sounds. In the above graph an R^2 of 42.08% was obtained with a third order polynomial equation of:

$$y = 0.0001x^3 - 0.0083x^2 + 0.1804x - 0.4569 \quad (1)$$

The linear regression equation obtained was (Figure 6):

$$y = 0.0152x + 0.4695 \quad (2)$$

and it yielded an R^2 of 18.84%, where $y = FEV_1/FVC$ and $x = HNR /ε:/$

The vowel sound /i:/ presented an R^2 value of 8.60% in correlating HNR with FEV_1/FVC ratio with a cubic polynomial regression. The equation obtained was:

$$y = -0.0004x^3 + 0.0211x^2 - 0.3282x + 2.2393 \quad (13)$$

The linear regression equation obtained was (Figure 7):

$$y = 0.0082x + 0.5634 \quad (14)$$

and it yielded an R^2 of 2.94%, where $y = FEV_1/FVC$ and $x = HNR /i:/$

The vowel sound /o:/ presented an R^2 value of 18.85% in correlating HNR with FEV_1/FVC ratio with a cubic polynomial regression. The

equation obtained was:

$$y = -0.0003x^3 + 0.016x^2 - 0.235x + 1.5459 \quad (5)$$

The linear regression equation obtained was (Figure 8):

$$y = 0.0121x + 0.4589 \quad (6)$$

and it yielded an R^2 of 6.73%, where $y = FEV_1/FVC$ and $x = HNR /o/$

The vowel sound /ɔ:/ presented an R^2 value of 20.75% in correlating HNR with FEV_1/FVC ratio with a cubic polynomial regression. The equation obtained was:

$$y = -0.0002x^3 + 0.0094x^2 - 0.0996x + 0.7913 \quad (7)$$

The linear regression equation obtained was (Figure 9):

$$y = 0.0137x + 0.4678 \quad (8)$$

and it yielded an R^2 of 9.56%, where $y = FEV_1/FVC$ and $x = HNR /ɔ:/$

The vowel sound /ɔ:/ presented an R^2 value of 6.53% in correlating HNR with FEV_1/FVC ratio with a cubic polynomial regression. The equation obtained was:

$$y = -0.0001x^3 + 0.0064x^2 - 0.0968x + 1.0298 \quad (9)$$

The linear regression equation obtained was (Figure 10):

$$y = 0.0074x + 0.5572 \quad (10)$$

and it yielded an R^2 of 4.17%, where $y = FEV_1/FVC$ and $x = HNR /u/$

The phrase "She sells" presented an R^2 value of 17.07% in correlating HNR with FEV_1/FVC ratio with a cubic polynomial regression. The equation obtained was:

$$y = -5 \times 10^{-6} x^3 - 0.002x^2 + 0.0769x + 0.0355 \quad (11)$$

The linear regression equation obtained was:

$$y = 0.0122x + 0.5043 \quad (12)$$

and it yielded an R^2 of 7.93%, where $y = FEV_1/FVC$ and $x = HNR$ "She sells"

As seen from the results above, the R^2 values obtained from the correlations between FEV_1/FVC and HNR were generally low. However, there were a few challenges during the data acquisition which could possibly have effects on the results. The challenges include:

i. Presence of background noise, as it was difficult to get a quiet place to take audio recordings. Thus, recordings were taken at relatively quiet locations but these were not without some level of significant noise, which is why most of the audio signals had to undergo noise removal.

ii. The recorder used was of multilateral and thus captured significant background noise even though it was placed close to patients' mouth. A unilateral recorder would have done better in this case.

iii. Some of the patients from which speech data was collected could not pronounce the vowels correctly.

iv. Some patients could not properly perform during the spirometry tests.

It is expected that addressing the aforementioned challenges would yield better results.

Conclusion

In this study, acoustic analysis was used to investigate the correlation between FEV_1/FVC ratio obtained from spirometry and Harmonics-to-Noise Ratio (HNR) of the vowels sounds /a:/, /e:/, /ɛ:/, /i:/, /o:/, /ɔ:/, /u:/ and phrase "she sells". It was found that the different sounds yielded different R^2 values. The results obtained have generally low R^2 values, the highest being 42.08% for the vowel /ɛ:/ with cubic polynomial regression. Challenges in speech and spirometry data collection could have greatly affected the results and thus it is recommended that any future work should address the challenges.

Acknowledgements

The authors would like to acknowledge the contribution of all who have in one way or the other aided in some aspects of this work especially, Dr. Audrey Forson of University of Ghana Medical School, Dr. Asomani of Chest Department, Korle-Bu, Ms. Beatrice Adom, Mr. Emmanuel Offei and Mr. Obed Korshie Dziku of the Department of Biomedical Engineering, University of Ghana. This work is fully funded by Office of Research and Innovation Development (ORID), University of Ghana. Grant Ref: ORID/ILG/019/05-13.

Informed Consent

Informed consent was obtained from all participants of the study before including them in the study.

References

1. Honda M (2003) Human speech production mechanisms. NTT Technical Review 1: 24-29.
2. Mohamed EE (2014) Voice changes in patients with chronic obstructive pulmonary disease. Egyptian Journal of Chest Diseases and Tuberculosis 63: 561-567.
3. Dixit VM, Sharma Y (2014) Voice Parameter Analysis for the disease detection. IOSR Journal of Electronics and Communication Engineering 9: 48-55.
4. Bousquet J, Khaltaev NG, Cruz AA (2007) Global Surveillance, Prevention and Control of Chronic Respiratory Diseases: A Comprehensive Approach. WHO.
5. Network GA (2014) The Global Asthma Report 2014. Auckland, New Zealand, p: 769.
6. Schlegelmilch RM, Kramme R (2011) Pulmonary function testing. In: Springer Handbook of Medical Technology, pp: 95-117.
7. Wanger J (2011) Forced Spirometry and Related Tests. In: Pulmonary function testing. Jones & Bartlett Publishers.
8. Sahebajami H, Gartside PS (1996) Pulmonary function in obese subjects with a normal FEV_1/FVC ratio. Chest 110: 1425-1429.
9. Swanney MP, Ruppel G, Enright PL, Pedersen OF, Crapo RO, et al. (2008) Using the lower limit of normal for the FEV_1/FVC ratio reduces the misclassification of airway obstruction. Thorax 63: 1046-1051.
10. Vollmer WM, Gislason P, Burney P, Enright PL, Gulsvik A, et al. (2009) Comparison of spirometry criteria for the diagnosis of COPD: results from the BOLD study. European Respiratory Journal 34: 588-597.
11. Farrús M, Hernando J, Ejarque P (2007) Jitter and shimmer measurements for speaker recognition. In: Eighth Annual Conference of the International Speech Communication Association.
12. Teixeira JP, Oliveira C, Lopes C (2013) Vocal acoustic analysis-jitter, shimmer and hnr parameters. Procedia Technology 9: 1112-1122.
13. Batra K, Bhasin S, Singh A (2015) Acoustic analysis of voice samples to differentiate healthy and asthmatic persons. International Journal of Engineering and Computer Science 4: 13161-13164.
14. Batra K, Bhasin S, Singh A (2015) Comparison of asthma and healthy persons using voice analysis. International Journal of Engineering Sciences & Research Technology 4: 928-932.
15. Yumoto E, Gould WJ, Baer T (1982) Harmonics-to-noise ratio as an index of the degree of hoarseness. The Journal of the Acoustical Society of America 71: 1544-1550.