Utilization of Low-Cost Sound Sensors with a built in Microphone as a Respiratory Pattern Sound Indicator and a Risk Mitigation Tool

In response to COVID-19

Gordana Laštovička-Medin
University of Montenegro
Podgorica, Montenegro
gordana.medin@gmail.com

Rajka Pejanović
University of Montenegro
Podgorica, Montenegro
p.r.rajka@gmail.com

Abstract—**Since the detection of pattern abnormalities may lead to not only the prevention of chronic respiratory diseases but also other diseases, many techniques have been developed in order to detect breathing and coughing patterns. To benefit from the cross-disciplinary studies we have decided to expose physics students to both: learning about sound using coughing as a targeted research topic and to develop a demo tool that is useful for building on exploratory skills and provides them with solid knowledge for future more advanced scientific research in biomedical engineering. A low-cost microphone sensor was tested for the purpose of understanding whether it can be used not only as a sound indicator but more broadly as a risk mitigation tool during a pandemic such as the current pandemic, COVID-19. The final goal of this long-term project is to build mathematical models aiding the identification of features from sound samples and to apply a classifier algorithm based on the machine learning technique at the final stage of research.**

Keywords- COVID-19, sound pattern, human generated cough and breathing, low-cost sound sensor, loudness

I. Introduction

COVID-19 is a respiratory infection caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. Prominent symptoms of COVID-19 include coughing and breathing difficulties. Cough sound analysis helps us to differentiate two similar sounds and to define the objective correlations with spirometry and clinical diagnosis [2], including Cough Peak Flow using cough sounds [3]. The auscultation of the respiratory system is another diagnostic technique and an inexpensive, noninvasive, safe, and easy-to-perform method [4].

The parameters such as frequency, intensity, and timbre of sound are of particular interest for the classification of respiratory diseases and are defined as follows. Pitch is the subjective perception of sound's frequency and depends on the frequency while amplitude of loudness is related to the energy of sound waves and is measured by the height of sound waves from the mean position; it is the subjective perception of amplitude. Quality or timbre is an important property of sound that differentiates two sounds with the same pitch and loudness. The fundamental frequency or primary frequency is the lowest frequency of a sound wave and it determines the pitch of the sound; the frequencies higher than the fundamental frequencies are called overtones while harmonics are overtones whose frequencies are whole number multiples of the fundamental frequency.

However, real-world sounds are not usually deterministic: they do not just have simple harmonics of the fundamental frequency. Instead, they also have unpredictable “inharmonic” frequencies that are not structured as noise. Thus having complete understanding of these measurable quantities and designing an experiment where those features are not lost when recorded and processed is crucial for further applications of machine learning techniques [5]. The issue is also how to deal with the research complexity without compromising the flexibility of techniques required for the extraction of sound features and still providing a comprehensive outcome that would not compress important information for the sake of data reduction. Towards this, this paper presents an early effort, mostly exploratory based, in building the capacity for such a complex and comprehensive task and towards creating a cough/respiratory sound database in Montenegro.

II. Experimental Procedure

A. Research question

The cough frequency is the most basic measure of coughing, but the objective study of cough signals has the potential to identify further features which may be clinically relevant and hence useful endpoints to study. Here we recall an early measurement (Fig. 1a) that used the audio tapes as research tools where the behavioral changes were monitored in order to extract untypical patterns over a longer time period [6]. Quantifying a cough was never an easy task, it is still not fully understood, and the symptoms are often incorrectly assigned. There is also no universally agreed unit of cough. The most intuitive way to quantify cough is to count the expulsive (first phase) of cough sounds (Fig. 1b). If long bursts (or peals) of expulsive cough sounds are present, then to identify each expulsive phase can be exceedingly difficult. The temporal patterns of coughing vary both in the short term (peals or epochs of coughs versus single coughs) and in the long term,

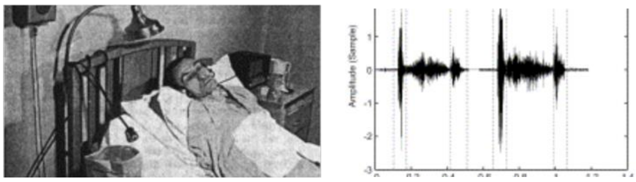

Figure 1. a) Early objective cough monitoring study with wall mounted microphone recording sounds from a hospital in-patient; b) Typical cough waveforms with expulsive phase, intermediate phase and voiced phase.

In another research project the respiratory sounds were recorded simultaneously with 16 microphones distributed over the thoracic surface [7]. Acoustic energy in three frequency bands, 150-300 Hz, 300-600 Hz and 600-1200 Hz, was analyzed during inspiration and expiration and results indicate that on average, inspiratory sounds are 10 to 11 dB louder than expiratory sounds at comparably flow rates. Findings that the acoustic shadow produced by the heart is more pronounced during inspiration support the concept that inspiratory sounds are produced predominantly in the periphery of the lung, while expiratory sounds are generated more centrally; lung sounds were also found to be significantly louder on the right side for the front part of the thorax, and on the left side for the back suggesting the importance of sensor positioning on the accuracy of the diagnostic procedure. [7]. Auscultation of the infant's chest reveals that their lung sounds seem to be different to those of adults since the normal lung sounds of newborn infants contain higher-frequency components than those of adults as a result of less filtering of the lung sound in infants [8]. Furthermore, accompanying airflow limitation or poor transmission of sounds was found for pulmonary emphysematous, indicating role of sound intensity as an important feature to be understood. From the point of view of the mechanism the sound is generated and then blocked over the transmission journey [9].

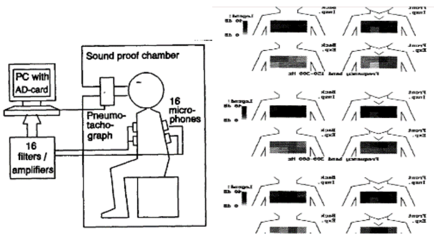

Figure 2. a) Sound loudness explored via microphones attached a 16 chest; b) The upper 4 diagrams represent the low frequency band, the middle 4 diagrams the intermediate frequencies and the lower 4 diagrams the high frequency band position. Microphones on the thoracic surface were arranged in a strictly geometric pattern to allow a meaningful comparison between different parts of the thorax [7].

It is important to also be aware that the testing of any new cough monitoring systems needs to be rigorous. Unfortunately, even the most recent medical systems suffer from a lack of sufficient accuracy while inconsistency in performance adds additional uncertainty in sound interpretation. It is also well known that any cough monitor will identify some cough events correctly (*true positives*), mistake non-cough events as coughs (*false positives*), miss some cough events (*false negatives*) and correctly ignore non-cough (*true negatives*).

All those problems and uncertainties in designing the data taking approach motivated us to set up our own investigation seeking for answers in long and short term. The research question is not only to experimentally test and demonstrate whether we can use a low-cost sound sensor module - MAX4466 Microphone Amplifier Module with an adjustable Gain Breakout Board for Arduino to detect sound patterns and utilize this tool for risk mitigation during COVID-19, or any respiratory related diseases, but also to learn more about the way the errors are overlooked and ignored (such as differentiation from ambient noise, differentiation from other patient sounds, especially speech, laughing sneezing and variability in the acoustics of cough sounds between individuals including transmission path of sound, positioning of producer of sound etc.).

B. Waveform patterns

Here we present some known waveform patterns that we will use to mimic "known/classified" sounds when requesting healthy volunteers to imitate the sound pattern. The different duration and pattern behavior of the three acoustic phases called the explosive, intermediate, and voiced (right) phases (introduced in the previous section) are clearly observed. Vesicular breath sounds (Fig. 3a) are heard across the lung surface. They are lower pitched, rustling sounds with higher intensity during inspiration. During expiration, sound intensity can quickly fade. Inspiration is normally 2-3 times the length of expiration [10],[12]. Bronchial breath sounds are tubular, hollow sounds which are heard when auscultating over the large airways (e.g. second and third intercostal spaces). They will be louder and higher pitched than vesicular breath sounds (Figure 3b). Figure 4a shows the wave form produced when one wheezes. As seen, wheezes can be high or low pitched. The waveform for crackles is displayed in Figure 4b. The pitch is lower than late inspiratory crackles. A patient's cough may decrease or clear these lung sounds.

Figure 3. Wave forms for: a) Vascular and b) bronchial breath.[10]

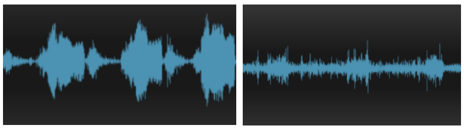

Figure 4. Waveform plot: a) Wheezes. b) Crackles [10]

C. *Set up*

The research tool was simple: it was based on Arduino and a microphone sensor: MAX4466 Microphone Amplifier Module, with electret microphones on it as shown in Figure 5. This breakout is applicable in projects such as voice changers, audio recording/sampling, and audio-reactive projects that use FFT. The Electret Mic Breakout translates amplitude (not volume) by capturing sound waves between two conducting plates (one a vibrating diaphragm and the other fixed) in the microphone and converting them into electrical waves. These electrical signals are then amplified and picked up by the microcontroller's ADC

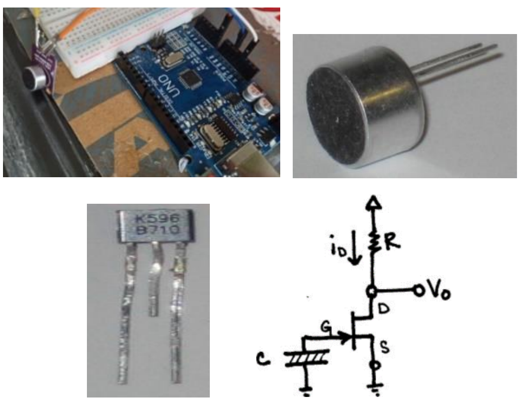

Figure 5. a) Sound sensor with microphone built in connected to the Arduino; b) Electret microphones c) Amplifier transistor (2SK596); d) Equivalent schematic of electret microphone.

Electret microphones: The amplifier consists of a single JFET transistor, with the gate connected to the pick-up plate, the source connected to ground, and the signal appearing on the drain. This is called a common-source configuration, as the source is connected to ground, which is common to all signals. The JFET in this electret microphone is a 2SK596, which is designed for low-noise applications. The electret maintains a fixed charge, and therefore maintains a voltage across the capacitor. For details we refer to [11].

Measuring Sound Levels: The Audio signal from the output of the amplifier is a varying voltage. To measure the sound level, we had to take multiple measurements to find the minimum and maximum extents or "peak to peak amplitude" of the signal. We have chosen a sample window of 50 milliseconds. That is sufficient to measure sound levels of frequencies as low as 20 Hz - the lower limit of human hearing. After finding the minimum and maximum samples, we computed the difference and converted it to volts and the output was printed on the serial monitor. We also did some experimental research with sounds at different volume levels to test how our average, min, max and span values respond. Adjustment to the the gain potentiometer was occasionally needed in order to utilize the max span for our sound levels while not overdoing it and not affecting the +/- sign of data.

D. *Calibration:*

The first step was to calibrate the sensor and to define the noise level (when no sound was produced). This procedure required additional adjustments of electronic components to provide optimal sensor performance. The optimal choice would be to have an acoustically isolated room which we were unable to obtain partially due to limits imposed by COVID-19 measures. Figure 5 shows the recorded noise level.

Figure 6. Sensor calibration: Noise level

E. *Row Data, Sampling, Response & Limits in sensor performance*

It is important to note that sound patterns are generated by volunteers, thus not clinically confirmed, and not taken from any approved medical data. For sampling the rule is to sample twice as fast as the maximum frequency (20 KHz) we want to capture. As we increase our knowledge about microphones and artefacts, we will learn more about the limits of its application in our research. Since, a JFET is used as the amplifier and because it has a high input resistance (30 MΩ or more) this means that almost no current is pulled off the electret capacitor. The reason for this is that if the amplifier had a lower input resistance, the low frequency response of the microphone would suffer. This is because the input stage acts like a high-pass filter, with the electret being the capacitor, and the input of the amplifier being the resistor, and larger values of R and C give lower cut-off frequencies. The main noise sources in this microphone are pick-up noise and transistor noise. Since the entire capsule is sealed and grounded, the pick-up noise is very low and usually not noticeable. The transistor noise, on the other hand, can be quite high, due to the high input resistance on the JFET. Typical values are around -120 dB to -110 dB, which may sound rather low, but the audio signal level is usually less than -40 dB, so it's only an 80 dB signal to noise ratio (SNR). This is a common issue with condenser microphones due to the high input resistances required. Regarding the sensor sensitivity we learnt that a smaller diaphragm would tend to give better high frequency and distortion characteristics, but will not be as loud, and therefore have worse SNR. A smaller diaphragm, as we found later in specifications, will also have a smaller capacitance, so its low frequency response will not be as good. Thus, picking a certain microphone for application must be rigidly tested.

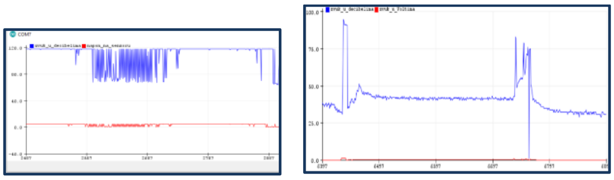

Figure 7. a) Raw data: after a period of silence a burst of cough is produced, and pattern repeated. b) Strong short burst of cough, then repeated after a prolonged time of silence. Some artefact added in the second burst.

Figure 7a show the relative sound level that the microphone picks up. The multiple samples were taken during a sample window which is set to 25 mS (50 mS = 20 Hz). The amplifier in this sound module is biasing the output at 1/2 of the Vcc used to power the board. The drawback here is that any audio (AC voltage) received and amplified will just randomly add or subtract from that 'center' value. Different scenarios for sound generation have been created. Fig. 7a) shows sound level picked up by the microphone after a certain period of silence was prolonged, then burst of coughs was produced, and this pattern further repeated. In Figures 7b) we slightly modified the scenario: a healthy male volunteer, after a period of being silent, (breathing quietly) generated strong but short coughs. We tried to mimic the crackles, but it becomes obvious that supervised training is required in order to obtain an optimal simulated pattern for crackles. However, the sensor was able promptly to record the change in sound pattern.

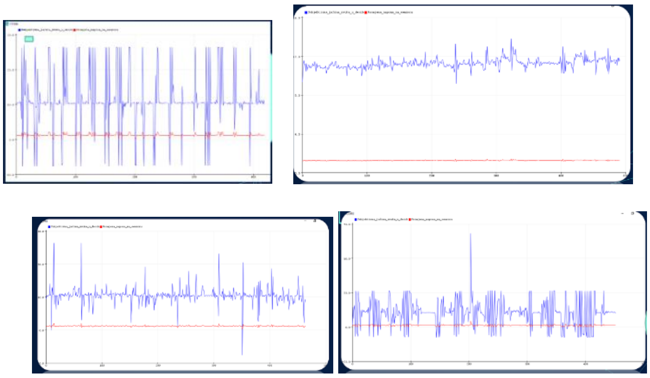

Figure 8. Raw data: a) Upper left image - Cough prolong over longer time, b) upper right image - monotonous breathing and c) bottom image - rapid breathing; d) singing; in the middle of singing a sharp clap was added (seen in image as spike)

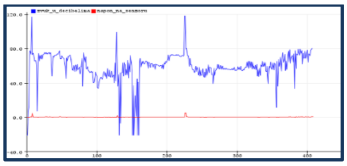

Figure 9. Randomly generated sound: Beathing, Coughing plus some artefacts.

The prolonged cough with the almost uniform intensity, the monotonous breathing, and the rapid breathing, are shown in, Figure 8a, 8b, and 8c, respectively. The variation of loudness while a person was singing is displayed in Fig 8d) while more randomly by human generated sound (in the pattern repetition and the amplitude variation) where breathing, short but strong coughing and for a noticeably short timing interval induced non-vocal artefact-sounds such by clapping, or screech (sharp spike) as presented in Figure 9.

F. Data Processing

The previous section was devoted to Data Acquisition. Approaching the pitch extraction is more challenging than the envelope extraction. The zero crossing technique was prone to errors because of harmonics causing additional zero-crossing. Techniques based on picks of the filtered sample might be a better option, but we struggled with the processed sample (only one we tested) since the fundamental frequency was week. The outcome of an FFT due to large "inharmonics" was difficult to interpret. However, we learnt a lot and we are currently working on gaining a better understanding of the techniques and the sampling. Regarding the envelope extraction technique, performed off-line, some issues had to be solved before we could proceed with the analysis. The cable that was used to connect the sensor to the computer had an issue and we solve it by providing better shielded. The noise that was constantly present in surrounding, was identified, recorded, then inverted in sign and added to the recorded sample. The artefacts affecting the duration of inspiration and expiration phases cause irreproducibility of the envelope extraction, and this is an internal feature of the signal envelope definition. An envelope of a signal is what is obtained through tracking successive ("connected") peak values and the technique can be described as follows: firstly, the signal is squared, then it is passed through a low pass 3rd order filter and in the final step the square root of the signal obtained from the previous step is calculated. Since, the extraction of the envelope is very useful for the classification of data by exploiting the features such as duration of inspiration and expiration phases giving us indications whether the analyzed data correspond to, for instance the crackled breathing where breath sounds are discontinuous and non-musical or to another respiratory disease, the control over all sources of uncertainty is crucial.

Many sound software for laptops can be used free of download but we suggest testing them first. Our first results on envelope extraction used data taken from a volunteer by requesting him to imitate crackled breathing with varying loudness and sampling was perform with frequency of 20 KZ (Figure 10). The expulsive phase and intermediate phase of cough is not distinguished well; from this we learnt that more work on filter algorithms has to be done. Some systematic shift was observed too (Fig 10b).

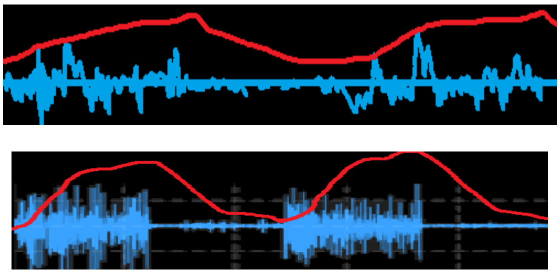

Figure 10. First insight into signal envelope extraction: a) upper image: volunteer mimic quite crackle breathing but sensor was occasionally lost the contact to computer; b) bottom image: Volunteer was asked to produce burst of cough and repat it. The systematic shift can be observed. More work on understanding the algorithm and outcome of analysis is undergoing.

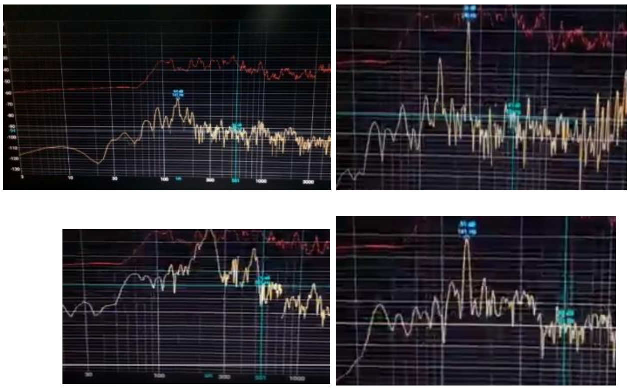

Figure 11. Power spectral density presented by blue line. The figure is the result of Fourier analysis and it represents the distribution of the intensity per frequency. The red line was left from previous measurement. The data filtering was not studied.

Few problems we identified. One issue is that we are converting an analogue input into a digital pin and this means that the triggering is not performed at a consistent voltage except if we connect the audio signal into a voltage comparator before it is converted into a digital signal through a digital pin; this way the pulse is always triggered at the same point or the same threshold voltage. Another thing is that the “envelope” only gives the information about the loudness, so one has to keep in mind that there is no frequency information in the extracted envelope. The real sounds are complex wave shapes, and it is hard to pick out when one cycle ends and the next starts, something that is relatively easy with the mathematically defined wave shapes such as squares. What makes it even harder is the fact that real sounds change their harmonic content and wave shape as the sound continues. One can learn a lot about envelope from analysing music notes played on string instruments and comparing the four envelope segments: the start of the sound is called the “attack” segment; after the loudest part of the sound, the fall to a steady “sustain” segment is called the “decay segment and when the sound ends, the fall from the sustain segment is called the “release” segment. Length of those segments can be used to differentiate different sources of sound; however, using only extracted envelope as feature for a certain type of cough seems to bring a lot of ambiguity. As excesses we play around the vibration, loudness and timbre of sound produced on cello and look at the change in the shape of sound envelope.

The quick FFT as a demo to students was performed without using an advanced data filtering approach, since the aim was to visualise change in frequency response as sound continue. An example of a power spectral density of a cough signal is shown in Fig. 11. The visible spikes present artefact. In future work the signal will be passed through the band pass filter to get rid of the low frequency noise. In this way we will eliminate most of the background noise present in surrounding including electronic sound of laptop ventilation. This will allow us to extract features such as the calculated length of a cough, the length between two hits in a cough attack and the total length of the cough sequence in processed data. It is also important to know that sound heard from mouth and those listened by practitioners from stethoscopes in auscultation differ from each other in the frequency content since some parts of body behave as low filters. This as an interesting topic to be further explored. We also found that some frequencies assigned as classifiers to some respiratory problems has overlapping with some musical notes (depends on what an octave is covered); so, the scope for a false positive is vast. We also noticed that some audio signals might have noise on them or harmonics in the waveform that cross the threshold voltage many times rapidly within the cycle we measured.

Since the MAX4466 is "optimized" for extremely low power, it might not be a good choice for comprehensive sound analysis, but it can be used as an experimental and learning tool. By visualizing what is lost and what information is gained using Fast Furrier Transformation with and without applying the data filtering one could learn more about changes in harmonic content and timbre (and resolution vs sample size which again indicate necessity for the good signal filtering due to huge processing time). On the other hand, the timbre that helps to distinguish two sounds with the same loudness depends on the relative strengths of the components of different frequencies and is mainly determined by the harmonic content of a sound and the dynamic characteristics of the sound such as vibrato and the attack-decay envelope of the sound. Thus, one has to be careful as to what extent the quality of original information or original content may be lost in data processing.

As a future exercise we might decide to use an area of microphone sensors adjusted to a certain frequency band using voltage control. Different microphone sensors will be explored too. Additionally, sensors such as a temperature, pressure sensor and an accelerometer attached to the chest would be an interesting and promising setup as well as an interesting new idea for those looking for more innovative and cross-disciplinary teaching approaches at the University.

III. CONCLUSION

In our study we explored the utilization of low-cost sound sensors with built-in microphones. We managed to demonstrate that our few-pound sound sensor is sensitive enough to distinguish different sound amplitudes and to precisely record the oscillation in sound pattern behavior.

Originally, we started this project as student lab exercise in response to COVID-19, and to test the proof of concept if a cheap microphone sensor could be useful (and to what accuracy level) in pandemic times (such as COVID-19) as a quick solution for risk mitigation (not applicable for clinical application before rigorous tests are performed). The presented analysis could be promising when it comes to differentiation of various sounds coming from the patient. The results are good starting point for further development of a solution that can be widely used for clinical purposes too.

The presented project at this level of development can be immediately utilized as a useful demo educational tool. The envelope extraction technique is also developed, and the first preliminary results are currently under inspection. More work on FFT as well as understanding the extracted envelope is undergoing. For this our priority is to develop a more comprehensive and robust filtering algorithm. After that we will process data to extract respiratory rate (RR). This would allow us to extract the features from the signal in form of vital signs important for COVID-19.

This research would further benefit from the inclusion of a targeted group of people with chronic respiratory problems. More data is needed to give a conclusion about the reliability of the sensor and its usage as a medical monitor device. However, preliminary tests clearly demonstrate that sensor is capable to follows variation in sound pattern indicating further that such a simple and low-cost set-up would be an efficient tool for sound mapping and sound monitoring of a certain environment and thus can be suitable aid for deaf people also during respiratory pandemics such as COVID.

We are also fond of sharing our experiences with readers. Open questions and issues we found have motivated us to work further on this topic. Filtering of the frequencies and harmonics during sound journey before it is captured by the sensor is also an interesting topic and research on it would give us more knowledge about lost information that might be useful in sound classification and the clarification of the uncertainties in its association to different respiratory diseases or other diseases. The published papers usually ignore the sources of uncertainty and we found it useful to open such questions.

ACKNOWLEDGMENT

We want to thank the MECO organizers for supporting our work by managing the continuation of the conference during the extremely challenging time of COVID-19 pandemic.

REFERENCES

[1] https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[2] Gowrisree Rudraraju and Sai Praveen Haranathc, “Cough sound analysis and objective correlation with spirometry and clinical diagnosis”, Informatics in Medicine Unlocked, Volume 19, 2020, 100319.

[3] Yasutaka Umayahara et al.,“Estimation of Cough Peak Flow Using Cough Sounds”, Sensors (Basel),. 2018 Jul 22;18(7):2381., doi: 10.3390/s18072381.

[4] Malay Sarkar, Irappa Madabhavi, Narasimhalu Niranjan, and Megha Dogra, “Auscultation of the respiratory system, Ann Thorac Med”. 2015 Jul-Sep; 10(3): 158–168., doi: 10.4103/1817-1737.160831

[5] Hans Pasterkamp, Steve S. Kraman and George R. Wodicka, “Respiratory Sounds Advances Beyond the Stethoscope”, American Journal of Respiratory and Critical Care Medicine, Volume 156, Issue 3, PubMed: 9310022, 1997

[6] Jaclyn Smith, “Ambulatory methods for recording cough”, Pulmonary Pharmacology & Therapeutics, Volume 20, Issue 4, August 2007, Pages 313-318

[7] Martin Kompis and George R. Wodicka, “Distribution of inspiratory and expiratory respiratory sound intensity on the surface of the human thorax”, Proceedings - 19th International Conference - IEEE/EMBS Oct. 30 - Nov. 2, 1997 Chicago, IL. USA

[8] J F Kanga, S S Kraman, “Comparison of the lung sound frequency spectra of infants and adults”, PMID: 3774386 DOI: 10.1002/ppul.1950020508

[9] Schreur HJ et al., “Lung sound intensity in patients with emphysema and in normal subjects at standardized airflows”, DOI: 10.1136/thx.47.9.674 PMID: 1440459 PMCID: PMC474797

[10] https://www.easyauscultation.com/vesicular-breath-sounds

[11] https:/www. openmusiclabs.com/learning/sensors/electret-microphones

[12] https://www.easyauscultation.com