Research Article

TimBre, cough based screening of Pulmonary Tuberculosis

Rahul Pathri [1], Shekhar Jha [2], Arpita Singh[3], Vaishnavi Reddy[3]], Samarth Tandon[3]

ABSTRACT: Shortness of breath & cough are the main reasons to visit a clinic world over. While Acute & Chronic cough are not mutually exclusive, Chronic cough is a manifestation of many diseases & it's "timbre" changes even with voluntarily elicited cough & is different for each pathological condition and also varies between the same subject. 95% of chronic cough entails Post Nasal Drip Syndrome (PNDS), Bronchial Asthma, Gastroesophageal Reflux Disorder (GERD), Chronic Bronchitis, Bronchiectasis & Angiotensin Converting Enzyme Inhibitor users (ACE) & Pulmonary Tuberculosis (TB) falls under the remaining 5% along with others.

The cough pattern is very distinctive in that it is Bitonal which may also mimic the cough sounds of a Bronchial compression [1]. Cough based analysis has been done in the past for TB which was confined to spectral analysis & clinical & demographics were not considered in the analysis which is extremely important in the post COVID-19 world. The Sustainable Development Goal (SDG3) goal of ending TB in India and elsewhere are fast approaching & there needs to be a screening mechanism to expedite active case finding which is taking a hit due to the pandemic. The global numbers of 10 Million TB cases is most likely to double if the necessary interventions are not made. The current solution is positioned as a screening tool which does not compete with a WHO approved product such as GeneXpert or TruNat, but instead is complimentary in nature in that, cases can be screened within minutes & channelled for a confirmatory diagnostic test in a developing high burden country (HBC) that has a need to accomplish SDG goals. Multi Drug Resistant TB & Human Immune Virus (HIV). Most importantly, avoid the infection of their kith & kin which exposes a vulnerable situation during a pandemic. This paper discusses findings, external hardware, consumables, clinical protocol, results of a multi-site double blinded clinical study at a large hospital of which a specific site related results are discussed at length for 500 subjects that achieved a sensitivity of 80% & a specificity of 92% paving way for it to be a screening solution that can be implemented at the last mile by a healthcare worker.

Keywords: Noise Reduction, Machine Learning, Spectral Analysis, Cough Screening, Point of Care, Non-Invasive, Real Time, Telemedicine

INTRODUCTION

Cough is Onomatopoeic in nature, meaning it is reproducible in that that the human ear can distinguish which gender it belongs to, dry or wet & is split into one to three sounds (phases). The Timbre of cough changes for each pathological state & even for voluntary cough sounds. Healthy cough sounds repeated multiple times usually have a similar spectral pattern but are slightly affected by temperature, humidity and pollution [1]

Chronic and Acute Cough are a resultant of various ailments and involve resonant sounds and often depend on the structure of the head, nose, neck diameter of a human being when they are finally heard by an experienced Physician. Here we mimic an experienced Pulmonologist or a Physician's judgement of the type of the cough that they have obtained after training their ears and brains over a period of time listening to a patient and also observing certain clinical characteristics such as Height/Weight and associated Body Mass Index (BMI), Appetite Patterns as an Ordinal Variable & other comorbidities to make a conclusion of it being TB or not and the Machine Learning Algorithms along with Spectral Analysis combine to be a replacement of the Physician than can be deployed anywhere by a healthcare worker. Most importantly, the cough is elicited and recorded onto a sophisticated Microphone Array with a after a deep breath while wearing a surgical mask which are followed as a part of the stringent protocol to avoid contaminating the environment with bacteria or virus and a following a general cough etiquette. The subjects cough was recorded indoors and outdoors with acceptable background noise that was filtered as a first step in data preparation or pre-processing. The cough elicited was voluntary in nature based on the fact that the Timbre does not change between voluntary or involuntary cough. A Timbre is determined by its Fundamental Frequency (F0), associated Harmonics and Overtones analogous to a stringed musical instrument or a percussion instrument. The harmonics being multiples of whole numbers of the Fundamental Frequency (F0) unlike the overtones. It is very intuitive to consider the F0 and Overtones as a pathological state that does not have a whole number equation. Dry cough with Dyspnoea is a classic symptom of latent or early heart disease & voluntary coughing is recommended during CPR to maintain sufficient blood flow to the brain & specially during Arrhythmia [1]. The benefits to explore cough as a screening/diagnostic tool are a plethora and with Telemedicine picking up steam, integrating this with other systems in the healthcare arena offer utmost value to the patient well being from both diagnostic and prognosis.

Motivation: SDG3 goals to eliminate and eradicate TB by 2025, Offering the solution which complements a confirmatory diagnostic test such as a GeneXpert or a TruNat to reach the said goal in high burden countries & most importantly present the findings of a third-party clinical trial.

Materials and methods:

Schematics of procedure is presented in figure 1 representing from cough sound collection, data analysis to reporting to concerned doctors.

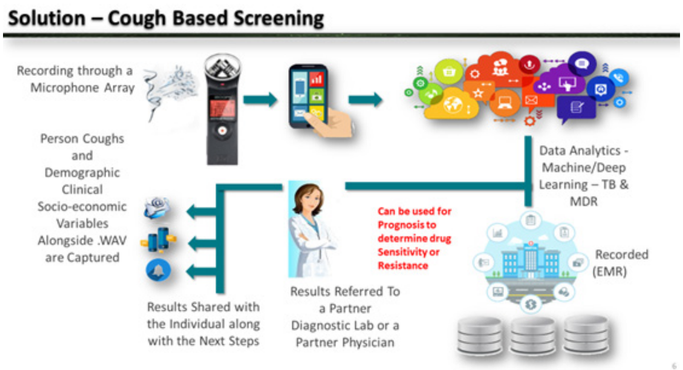


Fig. 1 Schematic representation of screening solution

Data Acquisition: The solution records a voluntary cough in a controlled environment using a third party Microphone Array (Zoom H1n) connected to a Mobile phone hosting an app that collects subject demographic, clinical information. A sampling rate of 44.1Khz with 16 bits was used as a function of the recording device. The same was also followed while building the TRAINING data set which included high resolution cough files arising out of zoom h1n and also native Android app that included low resolution cough files recorded with similar sampling and bit rates. A waiver was obtained while building the training set with a patient consent who were

Narayana Hrudayalaya Hospital for the Clinical Trials execution

Correspondence: Pathri R ,Narayana Hrudayalaya Hospital for the Clinical Trials execution and this project was funded by BIRAC/BIG, Axilor, Mumbai Angels.
Email: rahul@docturnal.com

mandated to wear a surgical mask during the cough recording. Clinical trials were conducted after obtaining the necessary ethics approval followed by subject consent. Underlying Machine Learning training data comprised of ailments not limited to Asthma, COPD, Bronchial Asthma, TB, MDR-TB, Pleural Effusion, Left Ventricular Disease (LVD), Empyema, Emphysema and Controls. The ground truth of annotating and labelling the data used case sheets from the Government Chest Hospital. This paved way for the ML Model to be an unbiased one. During the screening process, a strict protocol was followed. A new filter was used on the Microphone Array for each cough recording such that a subsequent subject cough recording did not risk any infection while the subject took a deep breath before the cough. The primary purpose of the filter was to eliminate extraneous noise. On top of this, the digital cough file was applied with NOISE reduction algorithm analogous to cocktail party algorithm written by the speech lab of IIITH.

Feature extraction: Using Fast Fourier Transform (FFT) [32] developed in Matlab & Python 3.0, a large number of summary variables/features were extracted from the audio signals in their frequency domain to study their predictive abilities. For the Clinical & Demographic variables, top performing ones were selected based on their usefulness in detecting and classifying Tuberculosis. Some of the significant features extracted for the study are described in Table 1.

Table 1 . Few of Fast Fourier Transform extracted features within the model with their equation and significance for prediction.

Mel-Frequency Cepstral Coefficient	$\log(S(w)) = \sum_{n=-\infty}^{\infty} c[n]\, e^{-jn\omega}$	Calculate the cepstral coefficients (c[n]) which represents the transformation of the power spectrum (s[w])& Randomness thereon
Energy Level	$Energy_{Level} = \sqrt{\frac{\sum_{n=0}^{N-1} x(n)^2}{N}}$	To calculate the Energy levels of the audio frame
Spectral Skewness Coefficient	$SSC = \left(\frac{E(x-\mu)}{\delta}\right)^3$	Measure of the asymmetry in the power spectrum about its mean (µ) was calculated using the spectral skewness coefficient and was used in understanding if the distribution of power contains density in the lower or higher frequency
Spectral Kurtosis Coefficient	$SKC = \left(\frac{E(x-\mu)}{\delta}\right)^4$	To measure the peak of the power spectrum
Spectral Centroid	$SC = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)} = \mu_1$	To calculate weighted mean of the spectrum which is also the representation of the centre of mass of the spectrum.
Spectral Flatness	$SF = \frac{e^{\left(\frac{1}{N}\sum_n \log(x(n))\right)}}{\frac{1}{N}\sum_n x(n)}$	To understand flatness of the spectrum by comparing the arithmetic with the geometric mean.

Classification Model and Valuation:

After the feature extraction part, RUSBoost a built–in MATLAB classifier [37] was used as a classification model to circumvent the Class Imbalance Problem. The TB to Non-TB split was 40:60. A Python implementation of the same was used to automate the result which had a concordance with the MATLAB implementation. RUSBoost uses a combination of random undersampling (RUS) and boosting. Random undersampling achieves a desired balance by continuously removing examples at random until the desired balance is reached. This improves the performance of weak classifiers by improving the accuracy of learning algorithm. Boosting in RUSBoost is an iterative process based on AdaBoost which iteratively constructs an ensemble of models[38] while adjusting the weights during each iteration. In this process the upcoming iterations correctly classifies the incorrect classifications in previous iteration& also assigns increased weights. The predictive models performance was evaluated by analysis of the area under curve (AUC) of the receiver operating characteristics (ROC) and with confusion matrices.

Results :

The training set at 40% TB and 60% non-TB, while covering an entire spectrum of Lung based ailments, still had a class imbalance and hence ENSEMBLE methods were used. The MATLAB R2018b classification learner model accuracy was at 86.4% with a Sensitivity and Specificity of 79% and 87% respectively with an Area Under Curve (AUC) of 0.89 for the binary classifier. A 10Fold cross validation was used to avoid over fitting as opposed to a partitioning of training and validation sets.

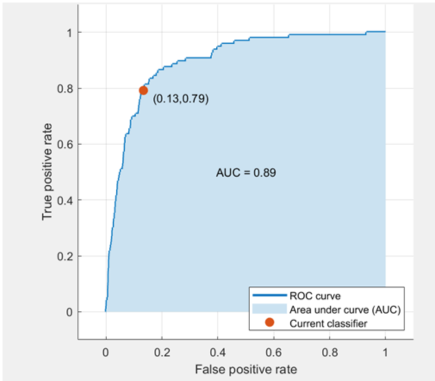


Fig. 2 : ROC curve of classification model.

The accuracy could be improved by eliminating Mobile recorded data in the training set while scoring Microphone Array recorded cough files & vice-versa. The ground truth around labels obtained from the case sheets was not always a molecular method but a combination of Chest X-Ray, Sputum Smear Microscopy and GeneXpert.

Non-Spectral Variables to derive BMI used a stadiometer and a digital weighing scale for height and weight that are typically carried by a healthcare worker reaching out to the last mile. Subject appetite patterns being an ordinal variable was obtained orally. Gender and Age were obtained from the case sheets of the site along with smoking and alcohol related data.

During the early phase of data harvesting, we have observed that the known TB cases consistently had a BMI < 17.5 and subjects complained of a loss of Appetite. Given the subjectivity around the level of appetite being low/medium/high, we decided to rely more on the BMI in the ML algorithm while converting the BMI into an Ordinal variable, revealed an Ordinal association between BMI and Positively screened cases. A Mantel-Haenszel Chi-Square Test revealed the same with a strong negative ordinal relationship between BMI category and Positive Cases as revealed by the Spearman Correlation Statistic.

The app was set to collect sound recording for around 20 seconds and cough for about 10 seconds after a deep inspiration. Healthy individuals themselves cannot cough for more than 10 seconds in reality. Subjects were asked to introduce themselves without disclosing any personally identifiable information (usually the first name) & cough onto the Microphone Array. While the app supported offline data collection, sound data was collected from the SDCard of the device where there were connectivity issues as opposed to streaming the sound file onto the HIPAA compliant cloud instance.

The actual clinical trial data is as follows for 474 subjects & while sensitivity of the model is pretty close to the actual findings, specificity of trial yielded a better value when compared to the model – 87% Vs 92%

Table.2 : Statistical findings of results.

Statistic	Value	95% CI
Sensitivity	80.00%	28.36% to 99.49%
Specificity	92.54%	89.77% to 94.75%
Positive Likelihood Ratio	10.72	6.24 to 18.43
Negative Likelihood Ratio	0.22	0.04 to 1.25
Disease prevalence (*)	1.05%	0.34% to 2.44%
Positive Predictive Value (*)	10.26%	6.23% to 16.42%
Negative Predictive Value (*)	99.77%	98.69% to 99.96%

Clinical Trial Results:

Primary Objectives of clinical trial:

1. To assess non-inferiority of sensitivity of TimBre software to not more than 10% of the sensitivity of standard screening modalities

2. Evaluate the diagnostic performance of TimBre with standard screening modalities as measured by sensitivity, specificity, PPV and NPV

The trial was divided into 3 groups of which group1 was a high prevalent Government chest hospital for 1000 subjects which is still work in progress & group3 at a private hospital for 100 subjects that had productive cough as a symptomatic condition is also work in progress that was disrupted due to the covid19 pandemic.

Group2 is what we are focused in this paper and comprised of 474 subjects recruited who were non-suspicious of TB. 435 subjects were screened as negative by TimBre and a confirmatory diagnostic procedure was not conducted for this segment of subjects but scheduled for the remaining two groups. The prevalence as seen from the table1 is at 1.05% and TimBre screened 39 as positive out of which 13 were lost to follow up and 12 were telephonically reached out for symptom screening. 7 telephonically counselled subjects who were located far away and could not visit are not discussed here given its subjectivity & also because they did not have any diagnosis on the Hospital Out Patient Department (OPD) Case Sheet. The remaining 20 are depicted in the table1 below that were subjected to various confirmatory diagnostic tests.

Table3. Clinical findings of patients with Timbre results.

patient_id	TimBre	method	result
msh38	Positive	cxr (chest x-ray)	nil
msh60	Positive	cxr	nil
msh132	Positive	followup	bronchial asthma
msh137	Positive	cxr	allergic bronchitis
msh146	Positive	cxr	bronchial asthma
msh157	Positive	cxr	nil
msh181	Positive	sputum for culture & sensitivity	nil
msh205	Positive	telephonic	bronchial asthma
msh229	Positive	sputum for culture/acid fast bacilli, cxr	negative
msh244	Positive	cxr	POSITIVE
msh256	Positive	telephonic	bronchial asthma
msh264	Positive	telephonic/cxr	negative
msh284	Positive	sputum for culture/acid fast bacilli	negative
msh314	Positive	ct scan chest	POSITIVE
msh326	Positive	cxr	POSITIVE
msh333	Positive	cxr	asthma
msh342	Positive	cxr, sputum acid fast bacilli	POSITIVE
msh360	Positive	telephonic	bronchial asthma
msh378	Positive	cxr	nil
msh389	Positive	sputum acid fast bacilli	negative

Four subjects screened as positive by TimBre were also confirmed by CXR (chest x-ray), CT-Scan Chest (computed tomography), CXR & CXR/Sputum AFB (acid fast bacilli) respectively. There were 5 nil diagnostics that did not result in any condition of the lungs. One of them was subjected to Sputum Culture & Sensitivity. 4 were negative when confirmed by Sputum Culture, Sputum AFB, CXR. We are unsure why culture was not considered as a gold standard and had to be added with a Sputum AFB in two cases (patient id: msh229 & msh284). 7 were confirmed as other lung conditions of which 6 were Bronchial Asthma and 1 was a case of Allergic Bronchitis. Nevertheless, all the 7 were comorbid conditions.

Table 4 : BMI and Appetite Pattern influence on the Model

patient_id	TimBre	method	result	BMI	appettite
msh229	Positive	sputum for culture/acid fast bacilli, cxr	negative	15.24158	Low
msh244	Positive	cxr	POSITIVE	13.71742	low
msh256	Positive	telephonic	bronchial asthma	17.63085	Medium
msh284	Positive	sputum for culture/acid fast bacilli	negative	28.65014	High
msh314	Positive	ct scan chest	POSITIVE	14.4795	Low
msh326	Positive	cxr	POSITIVE	15.05	Medium
msh333	Positive	cxr	asthma	15.427	Medium
msh342	Positive	cxr, sputum acid fast bacilli	POSITIVE	14.2432	Low
msh360	Positive	telephonic	bronchial asthma	17.00882	Medium
msh389	Positive	sputum acid fast bacilli	negative	17.79993	Medium

The Spearman Correlation Statistic based on the Mantel-Haenszel Chi-Square Test revealing ordinal association of BMI (ignoring appetite given its subjectivity), is quite obvious with the POSITIVE TB cases. It is also worthwhile noticing that Bronchial Asthma also has a decreased BMI. An outlier (false positive) here is patient_id msh284 that had a higher BMI that has been screened as POSITIVE by TimBre but is negative by a Sputum Culture test has to do with the spectral signature. A majority of the negatively screened subjects, numbering 435 had a BMI greater than 18 with a few outliers which are highly possible due to a human error while taking readings from the stadiometer and the digital weighing scale. Given the double blinded nature of the study, we could not control this aspect of data harvesting.

Explainable AI/ML:

A standardized parallel coordinate graph depicting the demarcation of TB and non-TB data for a spectral variable between 3500 to 4500 Hz in a single channel of a Mobile Phone. A total of 12 spectral feature sets were used across various frequency bands

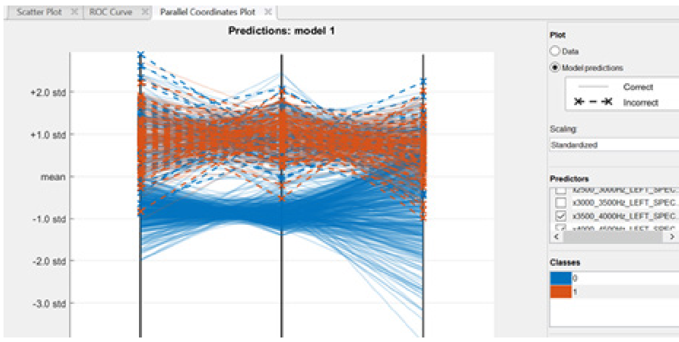


Figure 2 : Parallel Coordinate Graph for Spectral Centroid across relevant bands

Figure 3 a,b: Spectrograms, Waveforms & Enhanced Autocorrelation[39] with F0 from Audacity for the positively screened cough sounds after background noise removal for patient_id: msh244

Figure 3a

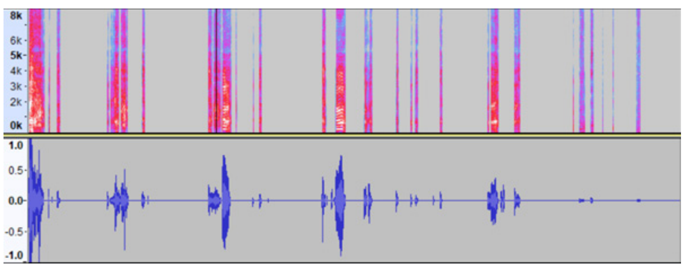


Figure 3b

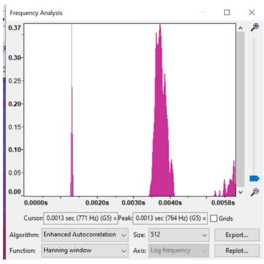


Figure 4a,b: Spectrograms, Waveforms & Enhanced Autocorrelation[39, 40] with F0 from Audacity for the negatively screened cough sounds after background noise removal for patient_id: msh342

Fig.4a

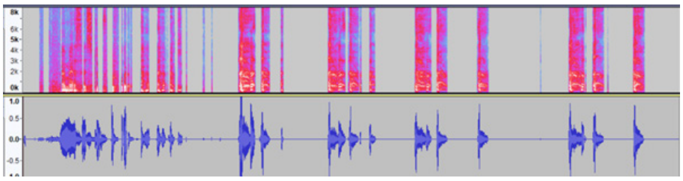
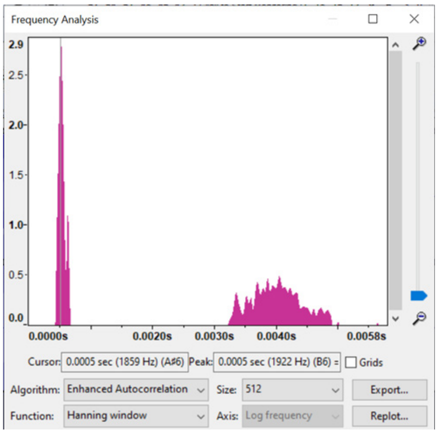


Figure 4b



Conclusion:

In conclusion, TimBre as a screening modality can be used for Mass & early detection while reducing the burden around infection. The opportunity for differential diagnosis [4] is tremendous wherein the false positives have been actually diagnosed as Bronchial Asthma for the most part and Bronchitis in some cases which still has an immense value for overall subject wellness around their lung health. Eliminating cases as non-TB is of humongous value given high specificity that can fit as a screening tool and become complementary in nature for a confirmatory diagnostic tool there by reducing the burden for such tests and focus on the right subjects that deserve a molecular test such as a GeneXpert and achieve the SDG goals in a much more seamless manner. Introducing this in the current workflow of TB diagnosis can be a big advantage in the public health space.

REFERENCES

1. 1) Cough: from Lab to Clinic by Ashima Anand(2007-12-01) – Book (Chapters 1,2,3,9)
2. 4) https://en.wikipedia.org/wiki/Differential_diagnosis
3. 32.Muqri R, Wilson J, Shakib J. A Taste of Python – Discrete and Fast Fourier Transforms; 2015
4. 37.Seiffert, C., T. Khoshgoftaar, J. Hulse, and A. Napolitano. RUSBoost: Improving classification performance when training data is skewed. 19th International Conference on Pattern Recognition, pp. 14, 2008.
5. 38. C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse and A. Napolitano, ”RUSBoost: A Hybrid Approach to Alleviating Class Imbalance,” in IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 40, no. 1, pp. 185-197, Jan. 2010
6. 39. Autocorrelation : https://en.wikipedia.org/wiki/Autocorrelation