Received: 3 September 2019 | Accepted: 27 December 2019
DOI: 10.1002/ppul.24633

ORIGINAL ARTICLE: NEONATAL LUNG DISEASE

Acoustic analysis of neonatal breath sounds using digital stethoscope technology

Lindsay Zhou MBBS[1,2] | Faezeh Marzbanrad PhD[3] | Ashwin Ramanathan MBBS[2] | Davood Fattahi MS[3] | Pramodkumar Pharande MBBS[1] | Atul Malhotra PhD[1,2]

[1]Monash Newborn, Monash Children's Hospital, Melbourne, Australia

[2]Department of Paediatrics, Monash University, Melbourne, Australia

[3]Department of Computer Systems and Electrical Engineering, Monash University, Melbourne, Australia

Correspondence
Atul Malhotra, Monash Children's Hospital, 246 Clayton Road, Clayton, Melbourne, VIC 3168, Australia.
Email: atul.malhotra@monash.edu

Funding information
Royal Australasian College of Physicians Research Fellowship; Advancing Women's Research Success Grant Program

Abstract

Background: There is no published literature regarding the use of the digital stethoscope (DS) and computerized breath sound analysis in neonates, despite neonates experiencing a high burden of respiratory disease. We aimed to determine if the DS could be used to study breath sounds of term and preterm neonates without respiratory disease, and detect a difference in acoustic characteristics between them.

Methods: A commercially available DS was used to record breath sounds of term and preterm neonates not receiving respiratory support between 24 and 48 hours after birth. Recordings were extracted, filtered, and computer analysis performed to obtain power spectra and mel frequency cepstral coefficient (MFCC) profiles.

Results: Recordings from 26 term and 26 preterm infants were obtained. The preterm cohort had an average gestational age (median and interquartile range) of 32 (31-33) weeks and term 39 (38-39) weeks. Birth weight (mean and SD) was 1767 (411) g for the preterm and 3456 (442) g for the term cohort. Power spectra demonstrated the greatest power in the low-frequency range of 100 to 250 Hz for both groups. There were significant differences ($P < .05$) in the average power at low (100-250 Hz), medium (250-500 Hz), high (500-1000 Hz), and very high (1000-2000 Hz) frequency bands. MFCC profiles also demonstrated significant differences between groups ($P < .05$).

Conclusion: It is feasible to use DS technology to analyze breath sounds in neonates. DS was able to determine significant differences between the acoustic characteristics of term and preterm infants breathing in room air. Further investigation of DS technology for neonatal breath sounds is warranted.

KEYWORDS
frequency, infant, lung, power, preterm, term

1 | INTRODUCTION

Physicians have been using auscultation of breath sounds as far back as the days of Hippocrates to yield diagnostic information, a skill further augmented by Laennec with his invention of the stethoscope in Paris 200 years ago. Since then, the stethoscope has become a ubiquitous aid to clinical examination and diagnosis in modern medicine.

Over the last several decades, recordings and acoustic analysis of breath sounds have been used to describe normal and pathological states beyond the scope of a traditional stethoscope.

Abbreviations: dB, decibels; DS, digital stethoscope; g, grams; Hz, hertz; MFCC, mel frequency cepstral coefficients.

The first lung sound spectrograms were presented in 1955,[1] followed by further study of normal lung sounds using spectral analysis in the 1980s.[2] Lung sound recordings were obtained with chest wall microphones and amplifiers to create a pneumo-phonogram and spectral analysis performed using the fast Fourier transform technique. This ability to analyze the acoustic qualities of recorded breath sounds has subsequently been tested in several adult and pediatric studies; investigating their relationship to normal physiology, pathology, and as a comparison and adjunct to the traditional stethoscope.[3,4] There are very few studies in medical literature documenting acoustic analysis of breath sounds in newborns, with only two neonatal studies found in a 2014 systematic review.[5] Pasterkamp et al[6] described the spectral characteristics of normal newborns in 1983, and Kanga et al[7] compared acoustic analysis of breath sounds from preterm infants who were self-ventilating in air, term infants, and adults. While this study included small numbers (six preterm babies), they were able to demonstrate differences between the three groups, with preterm infants demonstrating the highest average frequency, and adults the lowest. More recently, Ellington et al[8] conducted a computerized analysis of breath sounds in a larger cohort of normal children aged between 2 and 59 months, further characterizing breath sound profiles in the pediatric population.

The digital stethoscope (DS) has emerged more recently in pediatric medicine, with 25 studies of the technology identified in a recent systematic review,[9] including its use for monitoring of newborn heart rate during resuscitation,[10] and comparison with standard auscultation for detection of abnormal pediatric breath sounds.[11]

We aimed to determine if the DS could be used to study breath sounds of self-ventilating preterm and term neonates, and detect a difference in their acoustic characteristics. Given the very high burden of respiratory disease amongst preterm infants,[12] we determined this was the ideal cohort for which to investigate the technology further. The neonatal population also presents particular challenges in regard to studying breath sounds, including crying, mechanical background noise, small chest wall size for sensor contact, and irregular or periodic breathing.

We explored how the DS and computerized analysis could be used to obtain a spectral analysis of breath sounds in well preterm and term neonates breathing in room air, and potentially detect differences between the two groups related to size, lung and chest wall characteristics. This pilot study is the first step towards the use of this technology as a bedside tool for monitoring and potentially predictive use in neonatal medicine.

2 | METHODS

2.1 | Study design and participants

We conducted a prospective, single-center observational study of preterm and term neonatal breath sounds recorded using a commercially available DS (the Clinicloud Digital Stethoscope, Clinicloud Pty, Ltd, Melbourne, Australia). The study was conducted at Monash Newborn, Monash Children's Hospital, a tertiary-level neonatal unit in Melbourne, Australia, and was approved by the Monash Health Human Research Ethics Committee (HREA/18/MonH/471).

2.1.1 | Inclusion criteria

Preterm babies born at less than 36 weeks completed gestation, and full-term babies born between 37 and 42 weeks completed gestational age, who were not receiving respiratory support and between 24 and 48 hours old at the time of breath sound recording. Preterm babies were those being cared for in the Special Care Nursery and without respiratory support. Term babies were well infants on the maternity ward with no known medical conditions. We elected to study well infants in this study to establish normal values for digitally recorded breath sounds in term and preterm infants and to reduce interference from ventilation support devices and lung pathology in the first instance. Due to the hypothesis-generating, pilot design of the study, a sample size calculation was not performed.

2.1.2 | Exclusion criteria

Infants requiring any respiratory support, those with known congenital anomalies or genetic diagnosis, or parents unable to give consent.

Written, informed consent for breath sound recordings and demographic data collection was gained from the parents.

2.2 | Procedures

Breath sound recordings were made using the DS (Clinicloud Digital Stethoscope, Clinicloud Pty, Ltd) by placing it on the right anterior chest of the infant, and recording breath sounds for 60 seconds on to a commercially available smartphone software in mp3 format (VoiceRecorder, Tapmedia Ltd, London, UK). The right side of the chest was chosen to reduce interference from heart sounds, and anterior as the majority of infants are nursed supine and we aimed to minimize interference with the infants undergoing auscultation. Sound recordings were made when the infants were in a quiet, restful state either awake or asleep.

Recordings were made between 24 and 48 hours after the birth of the infant. This time frame was chosen to avoid potential interference from lung fluid clearance in the first 24 hours after birth affecting results, to avoid babies who may develop transient tachypnea of the newborn or respiratory distress syndrome, and after 48 hours most well full-term infants have already been discharged home.

2.3 | Analysis

Recordings were then downloaded, extracted and filtered using MATLAB (MathWorks, MA) by computer systems engineers. The engineers performing the sound feature extraction were blinded in regard to the origin of the recordings (preterm vs. term). A 10-second continuous segment of each recording containing breath periods was selected and filtered using a Butterworth bandpass filter at 100 to 2000 Hz to attenuate heart sounds and background noise. The sound recordings were analyzed in the frequency domain by Fast Fourier Transformation analysis, with the Welch method used for estimation of power spectral density. Analysis in the frequency domain allows for examination of the strength of a signal at a given frequency range, as opposed to the time domain which shows how a signal changes over time. Using this method transforms recorded sound to a two-dimensional numerical and graphical representation of sound wave power over frequency in Hz. Median power and interquartile ranges were obtained and plotted as power spectra in decibel (dB) over Hz.

Ratios of power at different frequency ranges designated lower 100 to 250 Hz, medium 250 to 500 Hz, high 500 to 1000 Hz, and very high 1000 to 2000 Hz were also calculated to obtain a proportion of power existing at a given frequency band. The frequency range of 100 to 2000 Hz was chosen based on previous studies of breath sounds in children yielding average power spectra in the range of 100 to 1000 Hz,[8] and we noted additional features on initial analysis beyond 1000 Hz so expanded the range to 2000 Hz.

Mel frequency cepstral coefficients (MFCCs) were also calculated for both groups by converting power spectra to the mel scale, which more closely models human hearing and represents perceived frequency vs actual measured frequency. A brief explanation of analysis in the frequency domain and MFCCs is available in Appendix A.

Data values acquired using MATLAB were then compared using IBM SPSS 25 (SPSS Inc, Chicago, IL). Shapiro-Wilk tests were used to determine the normality of each data set. For comparison between the normalized sound characteristic values of the two groups, Mann-Whitney U tests and paired sample *t* tests were used. For comparison of the two groups' demographic details, independent sample *t* tests and Mann-Whitney U tests were used to compare the normally and non-normally distributed features. Statistical significance was defined as a $P < .05$.

3 | RESULTS

3.1 | Patient recruitment

The parents of 58 infants were approached for inclusion in the study, and 55 infants' parents agreed to participate. About 55 recordings were made, and following exclusion of three due to poor sound quality (disconnection, majority nonbreath sounds), 26 preterm and 26 full-term babies were included for final analysis. Recordings took place over a 10-month period from September 2018 to May 2019 at Monash Newborn.

3.2 | Infant characteristics

The preterm cohort had an average gestational age (median and interquartile range) of 32 (31-33) weeks, and term cohort 39 (38-39) weeks. The average birth weight (mean and SD) was 1767 (411) g for the preterm group and 3456 (442) g for the term group. Demographic data is shown in Table 1.

3.3 | Breath sound analysis

Fifty-two breath sound recordings were downloaded, extracted, and filtered to obtain an average power spectrum for the two groups. Results were obtained in two formats: average power spectrum density in dB/Hz over Hz, and proportion of power existing at a low (100-250 Hz), medium (250-500 Hz), high (500-1000 Hz), and very high (1000-2000 Hz) frequency bands.

The average power spectrum for the two groups demonstrated the greatest power in the low-frequency range of 100 to 250 Hz, with an exponential decrease in power with increasing frequency. Average power spectrum density across the entire frequency range studied (100-2000 Hz) is demonstrated in Figure 1.

A significantly greater power was seen in the term infants compared with preterm infants over the medium and high-frequency range between 400 and 700 Hz, and a larger variance within the term cohort. At the very high-frequency range of 1000 to 2000 Hz, a significant and persistent separation in the two groups is again seen, but at this range, the preterm group demonstrates higher average power compared with the term (Figure 2), a reversal of the relationship at 400 to 700 Hz.

MFCCs were calculated for both groups over six frequency ranges, with results demonstrated in Table 2. Significant differences were obtained for MFCCs 2 (200-333 Hz), 3 (333-467 Hz), 5 (400-533 Hz), and 6 (467-600 Hz). When plotted against weight, there was a positive correlation between MFCC values and weight in grams (Figure 3).

TABLE 1 Demographic data

	Term infants (n = 26)	Preterm infants (n = 26)
Gestational age (wk)	39 (38-39)	32 (31-33)
Birth weight (g)	3456 (442)	1767 (411)
Male	15 (57%)	7 (27%)
Female	11 (43%)	19 (73%)
Mode of delivery		
Vaginal	12 (46%)	6 (23%)
Cesarean section	14 (54%)	20 (77%)
1 min Apgar score	9 (9-9)	8 (5-9)
5 min Apgar score	9 (9-9)	9 (8-9)
Multiple pregnancy	0 (0%)	18 (69%)

Note: Data represented as mean (SD), median (interquartile range), or total (percentage).

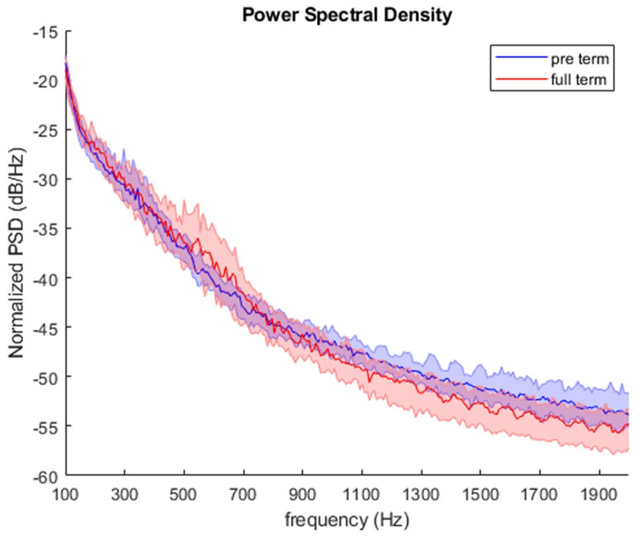

FIGURE 1 Average power spectral density in dB/Hz over 100-2000 Hz. Solid lines represent median, shaded areas interquartile ranges. dB, decibel; PSD, power spectral density [Color figure can be viewed at wileyonlinelibrary.com]

4 | DISCUSSION

This is the first study to analyze breath sounds in preterm and term neonates using DS technology and computerized sound analysis. Similar to previous studies in children and term infants,[6,8] the power spectra of breath sounds obtained were mostly contained within a 100 to 1000 Hz range, with most power in the low-frequency range of 100 to 250 Hz. We obtained statistically significant differences in the characteristics of breath sounds of self-ventilating preterm and term babies, with term infants' power spectra demonstrating greater average power at 400 to 700 Hz, but interestingly greater power in the preterm cohort from 750 Hz upwards through the very high-frequency range of 1000 to 2000 Hz.

It is encouraging to observe that this simple, noninvasive bedside technology was able to demonstrate differences in the acoustic qualities of breath sounds between preterm and term babies, given they were all self-ventilating in air, and such differences may not be detected by traditional auscultation. It is well-known that traditional auscultation is impaired by interobserver subjectivity and physician experience, particularly in the preterm cohort,[13] which is one of the reasons groups such as ours are investigating the use of the DS in pediatric medicine. We chose to study well infants first in this proof of concept paper using the DS and computerized analysis, allowing us to obtain a set of normal values without interference from ventilatory devices and significant lung pathology, before embarking on further study of infants requiring respiratory support.

We hypothesized that breath sounds of well, self-ventilating preterm and term babies may show differences in their power spectrum profiles, given previous studies demonstrated differences between age groups.[6-8] The differences observed between the two groups may relate to airway size, chest wall thickness, or a degree of subclinical hyaline membrane disease and immaturity in the preterm cohort.

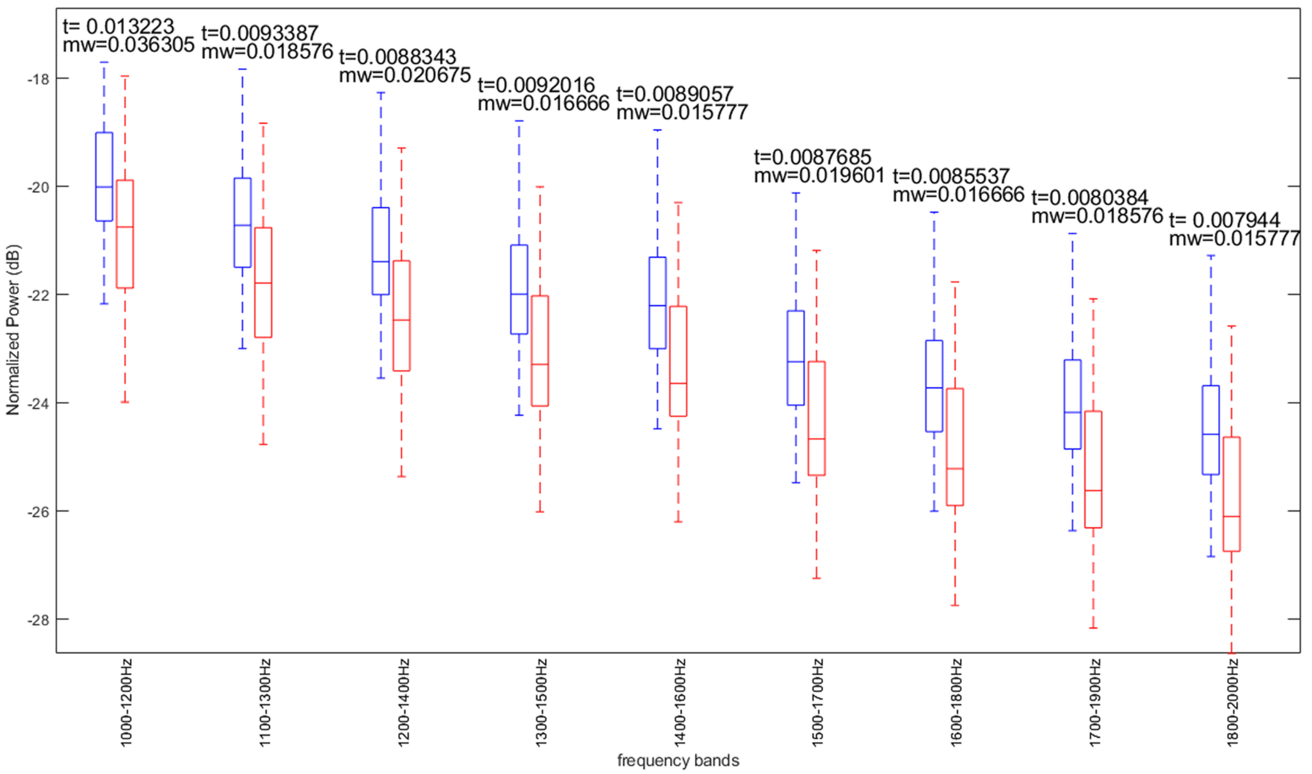

FIGURE 2 Normalized power at very high-frequency bands from 1000 to 2000 Hz. Blue box plots preterm, red full term. Box plots represent median-interquartile ranges and minimum/maximum values (shaded lines). *T* test and the Mann-Whitney U-test values sit above box plots. dB, decibel [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Mel frequency cepstral coefficient values (MFCCs)

MFCC values	Term infants (n = 26)	Preterm infants (n = 26)	Mean difference (95% CI)	*P* value
MFCC 1	−18.02 (1.00)	−17.91 (1.46)	0.103 (−0.594 to 0.799)	.76
MFCC 2	3.03 (0.29)	2.75 (0.21)	−0.272 (−0.414 to −0.131)	.0003
MFCC 3	0.77 (0.11)	0.55 (0.20)	−0.222 (−0.310 to −0.134)	<.0001
MFCC 4	0.56 (0.50-0.59)	0.55 (0.51-0.64)	−0.011 (−0.043 to 0.069)	.80
MFCC 5	0.23 (0.07)	0.31 (0.08)	0.080 (0.036 to 0.124)	.0007
MFCC 6	0.22 (0.06)	0.26 (0.06)	0.037 (0.005 to 0.069)	.02

Note: Data represented as mean (SD).
Abbreviation: CI, confidence interval.

Greater overall power was seen in the term cohort compared with preterm below 700 Hz. This may represent sound originating from a greater volume of aerated lung than in the preterm cohort, or greater power at lower frequencies generated by increased airflow in term infant lungs. Expiratory flow and respiratory compliance have been shown to be reduced in self-ventilating preterm infants compared with those born at full term,[14] which may be affecting the flow-mediated generation of breath sounds.

Previous studies have demonstrated that airflow affects lung sound intensity in adults,[15] and due to the size difference between the airways of preterm and term infants, airflow may be different. Malmberg et al[16] and Tabata et al[17] were able to document changes in acoustic qualities of breath sounds during bronchoconstriction challenges in asthmatic children, with greater power at high frequencies during bronchoconstriction, suggesting an association between narrower airways and higher frequency power spectrum profiles. This could explain the higher power seen in the preterm group above 700 Hz, who inherently have narrower airways than the term cohort.

Shirota[18] documented differences in acoustic qualities of lung sounds in term infants in the first 72 hours of life, and revealed a different spectral profile depending on the mode of delivery and changes over time, suggesting an influence of lung fluid clearance on breath sound profiles. It is well-known that preterm infants have a higher incidence of respiratory disease than those born at term,[19] and although the preterm infants in our cohort were asymptomatic, they may have experienced a degree of subclinical lung disease related to prematurity resulting in differences in breath sound characteristics. Elphick et al[4] demonstrated different acoustic spectra for infants with wheeze compared with those with "ruttles" (a palpable sound in the chest of babies), and Puder et al[20] were able to show detection of wheeze with computerized acoustic analysis using chest wall and tracheal sensors. These studies observed that lung disease states affect breath sound acoustic profiles, and it may be that both the preterm infants' size and potential subclinical lung disease has resulted in their different power spectrum profiles.

The MFCC values obtained are interesting, as they relate to how sound resonates from its source. The preterm babies were significantly smaller than the term babies, and therefore have narrower airways and thinner chest walls. We suggest that this has resulted in the difference in MFCC values, as the sound is traveling through smaller airways, chest cavity and thoracic wall than the term cohort, and this is supported by the positive correlation seen between MFCC values and infant weight.

Further study is required to determine the utility of this noninvasive technology in neonatology; this is particularly relevant for infants who have experienced extreme prematurity with a high burden of chronic respiratory disease, given the limitations of other diagnostic or predictive tools in respiratory medicine such as computed tomography (radiation) and respiratory function testing (impractical in small babies). It may be possible that electronically obtained sound features could predict bronchopulmonary dysplasia, and it has recently been demonstrated in a study of pediatric breath sounds that an artificial neural network was able to classify different types of respiratory sounds more efficiently than

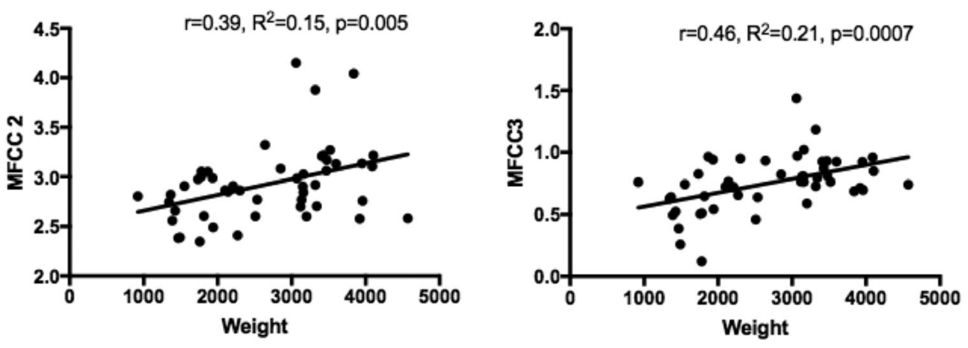

FIGURE 3 Mel frequency cepstral coefficient values 2 (200-333 Hz) and 3 (333-467 Hz) plotted against infant weight in grams

physicians.[21] This study is the first step towards the clinical utility of the DS and computerized breath sound analysis as a novel form of respiratory monitoring in neonatology, bedside breath sound analysis on a smartphone that may provide insights into lung pathology or future disease trajectory in real-time, as well as a potential role in telemedicine.

The main limitation of this study is its small size and pilot design, which limits the strength of the data obtained. However, despite the limited number of babies recorded, we were able to show feasibility in using this technology to obtain and describe preterm and term breath sound features.

5 | CONCLUSION

We were able to use the DS to easily record the breath sounds of preterm and term neonates breathing in room air, and use computerized analysis to describe their breath sound profiles. We were able to demonstrate significant differences between the sound characteristics of the two groups and suggest this relates to differences in lung, airway and chest wall size and characteristics. Further research using the DS in neonates is required, particularly for infants receiving respiratory support, investigation of changes to breath sounds over time, and of infants with different lung disease states. In an age of increasing overlap between medicine and technology, the combination of the DS, electronic data collection, and artificial intelligence could play an important role in neonatal respiratory medicine.

ACKNOWLEDGMENTS

We would like to thank the nursing and midwifery staff at Monash Medical Centre and Monash Newborn for their assistance during the recruitment period of this study. Dr Atul Malhotra is supported by a Royal Australasian College of Physicians Research Fellowship. Dr Faezeh Marzbanrad acknowledges the support of Advancing Women's Research Success Grant Program.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests. No author holds any interest in any of the commercially available stethoscope manufacturers.

ORCID

Atul Malhotra http://orcid.org/0000-0001-9664-4182

REFERENCES

1. McKusick VA, Jenkins JT, Webb GN. The acoustic basis of the chest examination; studies by means of sound spectrography. *Am Rev Tuberc*. 1955;72:12-34.
2. Gavriely N, Palti Y, Alroy G. Spectral characteristics of normal breath sounds. *J Appl Physiol Respir Environ Exerc Physiol*. 1981;50(2): 307-314.
3. Pasterkamp H, Kraman S, Wodicka G. Respiratory sounds – advances beyond the stethoscope. *Am J Respir Crit Care Med*. 1997;156(3): 974-987.
4. Elphick HE, Lancaster GA, Solis A, Majumdar A, Gupta R, Smyth RL. Validity and reliability of acoustic analysis of respiratory sounds in infants. *Arch Dis Child*. 2004;89:1059-1063.
5. Oliveria A, Marques A. Respiratory sounds in healthy people: a systematic review. *Respir Med*. 2014;108(4):550-570.
6. Pasterkamp H, Fenton R, Leahy F, Chernick V. Spectral analysis of breath sounds in normal newborn infants. *Med Instrum*. 1983;17(5): 355-357.
7. Kanga J, Kraman S. Comparison of the lung sound frequency spectra of infants and adults. *Pediatr Pulmonol*. 1986;2(5):292-295.
8. Ellington L, Emmanouilidou D, Elhilali M, et al. Developing a reference of normal lung sounds in healthy Peruvian children. *Lung*. 2014;192: 765-773.
9. Ramanathan A, Zhou L, Marzbanrad F, et al. Digital stethoscopes in paediatric medicine. *Acta Paediatr*, 108(5):814-822.
10. Kevat AC, Dawson J, Davis PG, Kamlin COF. Evaluation of a digital stethoscope and smart device technology for assessment of heart rate in the newborn infant. *Arch Dis Child Fetal Neonatal Ed*. 2015; 100(6):562-563.
11. Kevat AC, Kalirajah A, Roseby R. Digital stethoscopes compared to standard auscultation for detecting abnormal paediatric breath sounds. *Eur J Pediatr*. 2017;176:989-992.
12. Chow S, Creighton P, Chambers G, Lui K Report of the Australian New Zealand Neonatal Network; 2015. Australian New Zealand Neonatal Network, 2017. https://anznn.net/annualreports. Accessed August, 2019.
13. Jensen E, Panitch H, Feng R, Moore P, Schmidt B. Interobserver reliability of the respiratory physical examination in premature infants: a multicenter study. *J Pediatr*. 2016;178:87-92.
14. McEvoy C, Venigalla S, Schilling D, Clay N, Spitale P, Nguyen T. Respiratory function in healthy late preterm infants delivered at 33-36 weeks of gestation. *J Pediatr*. 2013;162(3):464-469.
15. Gavriely N, Cugell DW. Airflow effects on amplitude and spectral content of normal breath sounds. *J Appl Physiol*. 1996;80(1):5-13.
16. Malmberg L, Sorva R, Sovijärvi A. Frequency distribution of breath sounds as an indicator of bronchoconstriction during histamine challenge test in asthmatic children. *Pediatr Pulmonol*. 1994;18(3): 170-177.
17. Tabata H, Enseki H, Nukaga M, et al. Changes in breath sound spectrum during metacholine inhalation in children with asthma. *Respirology*. 2018;23(2):168-175.
18. Shirota K. Acoustic analysis of neonatal breath sounds: changes during the breath adaptation period. *J Jpn Bronchoesophagol Soc*. 1998;49(5):442-450.
19. Mahony A, Jain L. Respiratory disorders in moderately preterm, late preterm, and early term infants. *Clin Perinatol*. 2013;40(4):665-670.
20. Puder L, Wilitzki S, Buhrer C, Fischer H, Schmalisch G. Computerised wheeze detection in young infants: comparison of signals from tracheal and chest wall sensors. *Physiol Meas*. 2016;37:2170-2180.
21. Gryzwalski T, Piecuch M, Sjazek M, et al. Practical implementation of artificial intelligence algorithms in pulmonary auscultation examination. *Eur J Pediatr*. 2019;178(6):883-890.
22. Welch P. The use of fast Fourier transform for the estimation of power spectra: a method based on time-averaging over short, modified periodograms. *IEEE Trans Audio Electroacoustics*. 1967;15(2):70-73.
23. Solomon OM. *PSD computations using Welch's method (Power Spectral Density PSD)*. Technical Reports, United States doi:10.2172/5688766 available at: https://www.osti.gov/biblio/5688766-psd-computations-

using-welch-method-power-spectral-density-psd. Accessed December, 2019.

24. Davis S, Mermelstein P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans Acoustics Speech Signal Process*. 1980;28(4):357-366.
25. Muda L, Begam M, Elamvazuthi I. Voice recognition algorithisms using Mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. *J Comput*. 2010;2(3):138-143.
26. Lyons J. *Mel frequency cepstral coefficient (MFCC) tutorial*. Practical Cryptography 2012 (website), available at: http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/. Accessed December, 2019.

How to cite this article: Zhou L, Marzbanrad F, Ramanathan A, Fattahi D, Pharande P, Malhotra A. Acoustic analysis of neonatal breath sounds using digital stethoscope technology. *Pediatric Pulmonology*. 2020;1–7. https://doi.org/10.1002/ppul.24633

APPENDIX A: A BRIEF EXPANSION ON COMPUTERIZED ANALYSIS OF BREATH SOUNDS

Analysis in the frequency domain

This study used analysis in the frequency domain to obtain breath sound features, which requires conversion of the recordings which are taken in the time domain. Analysis in the frequency domain allows examination of signal power at a given frequency, as opposed to how it changes over time (as in the time domain).

There are various methods for the conversion of recorded sound for analysis in the frequency domain, and in this study, we used the Welch method. The Welch method reduces variance in nonhomogenous samples (eg, breath sounds mixed with heart sounds and background noise) by splitting it into small overlapping sequences of values called periodograms, calculated by discrete Fourier transform. These periodograms are then averaged to provide a power spectral density for the sample analyzed. This method was chosen to reduce the contamination of the power spectra obtained by nonbreath sounds, and is widely used for computerized sound analysis.

Mel frequency cepstral coefficients

The mel scale relates the perceived frequency of a sound to its actual frequency and is widely used in sound and voice recognition software to remove noise and focus on the most relevant portions of a recorded sound.

Mel frequency cepstral coefficients (MFCCs) are a well-known sound transformation used to extract the features which more closely model the human hearing system, by emphasizing the log-power spectrum in specifically scaled frequency bands.

This is because the human ear does not respond to sound in a linear fashion, rather on a logarithmic scale.

The procedure of calculating MFCCs can be summarized as follows, as used in this study:

- Frame the signal into short windows (in this study 30 millisecond frames with 20 millisecond overlap).
- For each frame calculate the power spectrum.
- Sum the energy over specific Mel scale frequency bands.
- Take the logarithm of the energies.
- Take the discrete cosine transform of the log energies.

The mel scales have a triangular shape amplifying the center frequency and attenuating the borders. The centers are chosen in the frequencies in which the human auditory system is more sensitive, and can then extract the recognition features. MFCCs relate to peak energy and resonance of sound and are affected by the system from which the sound originates, in the case of this study the lungs of the infants studied. (Additional References for Appendix: 22-26)