The unique heart sound signature of children with pulmonary artery hypertension

Mohamed Elgendi,[1,a] Prashant Bobhate,[2] Shreepal Jain,[2] Long Guo,[2] Shine Kumar,[2] Jennifer Rutledge,[2,3] Yashu Coe,[2,3] Roger Zemp,[4] Dale Schuurmans,[1] Ian Adatia[2,3]

[1]Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada; [2]Department of Pediatrics and Pediatric Pulmonary Hypertension Service, Stollery Children's Hospital, University of Alberta, Edmonton, Alberta, Canada; [3]Mazankowski Alberta Heart Institute, Edmonton, Alberta, Canada; [4]School of Biomedical Engineering, University of Alberta, Edmonton, Alberta, Canada; [a]Current address: Electrical and Computer Engineering in Medicine Group, University of British Columbia, and British Columbia Children's Hospital, Vancouver, British Columbia, Canada

Abstract: We hypothesized that vibrations created by the pulmonary circulation would create sound like the vocal cords during speech and that subjects with pulmonary artery hypertension (PAH) might have a unique sound signature. We recorded heart sounds at the cardiac apex and the second left intercostal space (2LICS), using a digital stethoscope, from 27 subjects (12 males) with a median age of 7 years (range: 3 months–19 years) undergoing simultaneous cardiac catheterization. Thirteen subjects had mean pulmonary artery pressure (mPAp) < 25 mmHg (range: 8–24 mmHg). Fourteen subjects had mPAp ≥ 25 mmHg (range: 25–97 mmHg). We extracted the relative power of the frequency band, the entropy, and the energy of the sinusoid formants from the heart sounds. We applied linear discriminant analysis with leave-one-out cross validation to differentiate children with and without PAH. The significance of the results was determined with a *t* test and a rank-sum test. The entropy of the first sinusoid formant contained within an optimized window length of 2 seconds of the heart sounds recorded at the 2LICS was significantly lower in subjects with mPAp ≥ 25 mmHg relative to subjects with mPAp < 25 mmHg, with a sensitivity of 93% and specificity of 92%. The reduced entropy of the first sinusoid formant of the heart sounds in children with PAH suggests the existence of an organized pattern. The analysis of this pattern revealed a unique sound signature, which could be applied to a noninvasive method to diagnose PAH.

Keywords: pulmonary hypertension, congenital heart disease, auscultation, machine learning, language recognition.

Untreated pulmonary artery hypertension (PAH) is a progressive, fatal disease.[1] It complicates many conditions and may affect up to 100 million people worldwide.[2,3] PAH is difficult to diagnose because symptoms appear late in the disease course and the findings on clinical examination are missed easily.

The finding on auscultation of a loud pulmonary component of the second heart sound (S2) in PAH has led to the exploration of phonocardiographic associations between S2 and pulmonary artery pressure (PAp) in the time domain.[4-10] However, precise demarcation, timing, and segmentation of the components of S2 remain challenging.[7-9,11-13]

We have explored instead quantitative information in the frequency domain of heart sounds that distinguish between subjects with and without PAH.[14] The relative power of the frequencies between 21 and 22 Hz of the heart sounds recorded at the second left intercostal space (2LICS) was significantly reduced in subjects with PAH.[14] However, there was a 22% error in detecting PAH. Therefore, by investigating further the recordings in these same subjects,

we sought to explore other features of the heart sounds in this specific frequency domain that might contain a unique feature that would identify subjects with PAH.

Normal speech patterns have a unique signature related to vocal cord vibration, which can be used, for example, to recognize a speaker as male or female. We hypothesized that vibrations created by the movement of the pulmonary valve leaflets or pulmonary artery would create sound in a manner similar to that of vocal cords during speech and that subjects with PAH might present a unique sound signature. The frequency resonance of sound is called a "formant." Formants are concentrations of energy that are prominent in a sound spectrogram and collectively constitute the identifying frequency spectrum for a sound produced by speech.

The relative positioning of the first and second formants is usually sufficiently unique to distinguish speech sounds, thus imparting a special quality, or timbre. Therefore, we investigated whether the energy and entropy of the first formant of recorded heart sounds could distinguish subjects with PAH. We hypothesized that vibra-

Address correspondence to Dr. Ian Adatia, Room 3A3.42, Stollery Children's Hospital, 8440 112 Street NW, Edmonton, Alberta T6G 3B7, Canada. E-mail: iadatia@ualberta.ca.

tions created by the pulmonary circulation would create sound, like the vocal cords during speech, and that subjects with PAH might have a unique sound signature.

METHODS

The University of Alberta Research Ethics Board approved the study. All subjects or their parents gave informed and written consent to participate in the study. Informed assent was obtained from children who were developmentally able to do so.

Clinical-data collection

We included consecutive subjects who were undergoing right heart catheterization as a requirement for managing their underlying cardiac condition. We excluded subjects with an abnormal or prosthetic valve.

The direct measurement of PAp, collected simultaneously with the heart sounds, was obtained with fluid-filled catheters in a standard manner. The heart sounds were recorded with a 3M Littmann 3200 electronic stethoscope (3M, Copenhagen, Denmark), which

works in conjunction with Zargis Cardioscan software (Zargis Medical, Princeton, NJ) to store recorded heart sounds in *.wav mono audio format. Heart sound recordings were obtained over 20 seconds with a sampling frequency of 4,000 Hz. We recorded the heart sounds sequentially at the 2LICS and over the cardiac apical impulse. For signal analysis and optimization, MATLAB 2010b (MathWorks, Natick, MA) was used.

Definition of PAH

PAH, in adults and children, is defined as mean PAp (mPAp) $\geq$ 25 mmHg and pulmonary artery wedge pressure (PAWp) or left atrial pressure (LAp) $\leq$ 15 mmHg measured at cardiac catheterization in subjects at rest.[15-17]

Definition of entropy

We defined entropy as a measure of the disorder of the heart sound pattern. A lower entropy value suggests the existence of an organized heart sound pattern, while a higher entropy value indicates more disorder.

Table 1. Summary of the demographic and hemodynamic data of all subjects

	PAp $\geq$ 25 mmHg		PAp < 25 mmHg	
	Median[a]	Range	Median[a]	Range
Age, years	8.5	0.8–15	3	0.3–19
Height, m	1.3	0.6–1.6	1	0.5–1.8
Weight, kg	31	6–78	18	4–59
BSA, m^2	1.1	0.3–1.8	0.7	0.2–1.7
BMI, kg/m^2	18	14–33	17	12–24
Sex	7M : 7F		5M : 8F	
Mean PAp, mmHg	54	25–97	15	8–24
Systolic PAp, mmHg	69	14–140	25	25–34
Diastolic PAp, mmHg	3	12–66	9	4–17
Mean LAp, mmHg	7	2–14	7	1–11
PVRI, WU · m^2	10	5–27	2.3	0.6–6
QPI, L/min/m^2	3.3	2.6–5.5	3.6	2.3–14
Mean BP, mmHg	70	48–96	60	42–116
Systolic BP, mmHg	94	44–122	92	63–110
Diastolic BP, mmHg	53	34–73	46	32–96
Mean RAp, mmHg	4	1–8	2	1–11
Heart rate, beats/min	84	63–130	100	70–134
QRS duration, ms	86	62–132	103	77–147
PR interval, ms	91	66–124	97	76–103

Note: PAp was measured during auscultation. PVRI is calculated from mean PAp measured at the time of thermodilution or oxygen consumption measurement and oximetry. BMI: body mass index; BSA: body surface area; BP: systemic blood pressure; LAp: left atrial pressure; PAp: pulmonary artery pressure; PVRI: pulmonary vascular resistance index; QPI: pulmonary blood flow index; RAp: right atrial pressure; WU · m^2: Wood unit.

[a] For sex, the ratio is reported.

Heart sound analysis

We classified the subjects and their heart sound recordings into two groups based on whether their mPAp was ≥25 mmHg or <25 mmHg. All subjects had PAWp < 15 mmHg. We extracted three spectral features: the relative power, energy, and entropy of the first 4 sinusoids of the heart sound frequency bands. We undertook separability tests to discover which recording site (the cardiac apex or the 2LICS) was more informative for the diagnosis of PAH. We applied linear discriminant analysis (LDA) to the most informative feature, with the aim of distinguishing subjects with PAH.

Feature extraction

A main part of our data analysis was the extraction of a feature from the heart sounds that provided the highest prediction rate for PAH. As discussed above, we began this process by collecting 2 heart sound recordings from 2 separate sites (the 2LICS and the apex) for each subject. Then we determined which site was the more informative site for PAH prediction. Once this site was identified, we extracted features that identified PAH. We selected the feature that provided the highest prediction for PAH from among all of the identified features.

The features that were extracted from the optimal site were relative power of the frequency band, energy, and entropy. Below, we expand on the detailed descriptions of each feature.

Relative power of the frequency band

The relative power of a frequency band is obtained by dividing the power of the band by the total power. However, in this investigation the relative power was calculated by dividing the power of the 21–22-Hz frequency band by the power of the 1–80-Hz frequency band, as suggested by our previous work.[14]

Sine wave heart sound replicas

Formants (unique heart sound signatures) were extracted by means of sine wave replicas, which distill the sound patterns down to key elements by removing extraneous noise. The audio track of each heart sound recording was transformed into sine wave replicas, as in speech analysis.[18] These sine wave replicas were transformed by tracking the frequencies and amplitudes of the first 4 formants as they varied over time. The acoustic measurements were obtained in a 2-step process. First, each sound file was resampled to 8 kHz. The resampled heart sound recording was then broken into 32-millisecond windows. Each window was subjected to an eighth-order linear-predictive-coding (LPC) analysis. LPC finds the coefficients of an eighth-order linear predictor (finite-impulse-response filter) that predicts the current value of the heart sound segment on the basis of past samples. The 4 coefficients with the highest magnitudes were converted to frequencies and stored in a data file. Each heart sound recording thus had an associated data file with 8 parameters (4 frequencies and 4 associated amplitudes) measured in each 32-millisecond window. This window captured information sufficient to track the change of the major formants in the original sound file over time. These data were submitted to a synthesis routine developed by Ellis,[19] which produced 4 sinusoidal tones that varied over time. We calculated the spectrogram (short-time Fourier transform) for each sinusoid, which we refer to as S in the equations below.

Energy and entropy

Energy.   We calculated the energy of a sinusoid as the power of the spectrogram: energy $= \Sigma_{i=1}^{L} S_i^2$, where L is the length of the processed heart sound segment.

Entropy.   We calculated the entropy of a sinusoid as the power of the log-transformed spectrogram: entropy $= \Sigma_{i=1}^{L} S_i \log(S_i)$, where L is the length of the processed heart sound segment.

LDA

We applied LDA to classify patients as either having PAH or normal, on the basis of the entropy of the first sinusoid formant of the heart sound. In studies where the sample size is small and cross validation is needed—such as our study—leave-one-out (LOO) is the only available method to estimate how accurately a predictive model will perform in a real practice setting. We assessed the classification performance with LDA through LOO cross validation to

Table 2. Comparison of clinical and hemodynamic data between subjects with pulmonary artery hypertension (mPAp ≥ 25 mmHg) and subjects with normal pulmonary artery pressures (mPAp < 25 mmHg)

Clinical and hemodynamic variables	P value
Age	0.5
Height	0.4
Weight	0.2
Body surface area	0.3
Body mass index	0.4
Systolic PAp	<0.001*
Diastolic PAp	<0.001*
Mean PAp	<0.001*
PVRI	<0.001*
Pulmonary blood flow index	0.9
Mean LAp	0.9
Mean RAp	0.5
Systolic blood pressure	0.2
Diastolic blood pressure	0.2
Mean blood pressure	0.1
Heart rate	0.2
ECG QRS duration in lead V1	0.02*
ECG PR interval in lead 2	0.4

Note: ECG: electrocardiogram; LAp: left atrial pressure; PAp: pulmonary artery pressure; PVRI: pulmonary vascular resistance index; RAp: right atrial pressure.
    * P < 0.05.

determine how the results of our statistical analysis would generalize to an independent heart sound data set.

Throughout the LOO process, each patient provided one case. Each training set was constructed by taking all cases except one, which was held out as a disjoint training set. For each training set, an LDA classifier was produced whose classification accuracy was determined on the single held-out test case. The average accuracy over all n splits was determined.

Nonstationarity

Heart sounds are noisy and highly nonstationary. If a heart sound signal was stationary, one could use the entire signal to calculate the spectral features (relative power of the frequency band and energy and entropy of the first sinusoid formant). However, since they are nonstationary, such features can vary over time, making it no longer meaningful to estimate the features from the entire 20-second duration of the signal. Thus, to accommodate potential nonstationarity, we conducted a search over segments of the heart sound recordings to identify an appropriate window length, L, that would allow accurate classification. Window lengths between 1 and 20 seconds were systematically tested. For each window length, the spectral feature was calculated and then averaged over all disjoint segments of length L within the duration of the 20-second heart sound recording.

Statistical tests

We calculated each spectral feature (relative power of the frequency band or energy or entropy of the first sinusoid formant) for each heart sound recording. As we had 27 subjects, each spectral-feature set contained 27 values (13 values from subjects with mPAp < 25 mmHg and 14 values from subjects with mPAP ≥ 25 mmHg).

To demonstrate significance of the mean and median of the samples within each spectral-feature set, we compared the values within each spectral-feature set by applying two tests: the 2-sample t test to compare the means and the 2-sided Wilcoxon-Mann-Whitney test (rank-sum test) to compare the medians. After calculating the P values returned from the t and rank-sum tests, the overall P value was calculated as $(P_{t\text{test}} + P_{\text{rank-sum}})/2$, with $P \leq 0.05$ considered significant. We repeated this procedure over a range of window lengths of heart sounds (varying from 1 to 20 seconds; see "Nonstationarity" above) for the three spectral features to determine the most informative window length for each feature.

Since we considered three different features and many different window lengths settings simultaneously, it is likely that a few P values are small merely as a result of stochastic fluctuations rather than as a result of systematic difference between subjects with mPAP < 25 mmHg and those with mPAP ≥ 25 mmHg. As a consequence, the P values have to be appropriately corrected. One may try to

Table 3. Window size analysis of three spectral features of the heart sounds at the second left intercostal space (2LICS): relative power of the frequency band 21–22 Hz, energy of first sinusoid formant, and entropy of first sinusoid formant

Window length, s	Relative power		Energy (first sinusoid)		Entropy (first sinusoid)	
	2LICS	Apex	2LICS	Apex	2LICS	Apex
1	4.24E−01	8.45E−01	1.10E−04	1.22E−01	4.15E−05	8.73E−02
2	1.41E−01	8.20E−01	8.52E−05	1.20E−01	**3.23E−05**	9.71E−02
3	6.86E−02	7.75E−01	7.05E−05	1.26E−01	3.30E−05	1.44E−01
4	4.29E−02	8.10E−01	9.94E−05	1.36E−01	5.14E−05	1.17E−01
5	3.19E−02	7.90E−01	**6.01E−05**	1.34E−01	5.02E−05	8.71E−02
6	3.20E−02	7.95E−01	6.28E−05	1.24E−01	4.21E−05	9.35E−02
7	3.13E−02	8.10E−01	6.53E−05	1.17E−01	4.38E−05	1.34E−01
8	2.86E−02	8.15E−01	6.54E−05	1.17E−01	5.20E−05	1.27E−01
9	2.44E−02	8.05E−01	6.48E−05	1.17E−01	4.15E−05	1.11E−01
10	2.39E−02	8.20E−01	6.40E−05	1.18E−01	4.98E−05	1.01E−01
11	**2.27E−02**	8.25E−01	6.51E−05	1.09E−01	5.15E−05	1.01E−01
12	2.27E−02	8.20E−01	6.47E−05	1.13E−01	5.22E−05	9.37E−02
13	2.31E−02	8.15E−01	7.47E−05	1.20E−01	5.16E−05	9.00E−02
14	2.43E−02	8.30E−01	7.31E−05	1.18E−01	6.33E−05	6.79E−02
15	2.44E−02	8.05E−01	7.10E−05	1.20E−01	7.91E−05	8.81E−02
16	2.41E−02	8.00E−01	6.12E−05	1.30E−01	9.96E−05	1.20E−01
17	2.71E−02	7.80E−01	5.99E−05	1.36E−01	8.82E−05	1.09E−01
18	2.80E−02	7.65E−01	8.36E−05	1.22E−01	1.13E−04	1.02E−01
19	2.91E−02	7.60E−01	1.04E−04	1.08E−01	6.66E−05	9.20E−02

Note: Data are overall P value $((P_{t\text{test}} + P_{\text{rank-sum}})/2)$. Boldface indicates the lowest P value for each feature.

This content downloaded from 096.052.061.159 on August 01, 2016 14:20:25 PM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

control the probability that a false positive occurs by applying Bonferroni postcorrection.[20] Since we were dealing with multiple comparison tests (114 tests in total), it was preferable to control for the false-discovery (false-positive) rate. Therefore, we used the Holm-Bonferroni method, as it controls the false-positive rate and offers a simple test uniformly more powerful than the Bonferroni correction.[21]

Two statistical measures were used for the output of the LDA analysis: sensitivity, which was calculated from the formula TP/(TP + FN), and specificity, which was calculated from the formula TN/(TN + FP), where TP is the number of true positives (PAH subjects detected as PAH subjects), FN is the number of false negatives (PAH subjects detected as normal-PAp subjects), TN is the number of true negatives (normal-PAp subjects detected as normal-PAp subjects), and FP is the number of false positives (normal-PAp subjects detected as PAH subjects).

RESULTS

We collected recordings from 27 subjects (12 males and 15 females) with a median age of 7 years (range: 3 months–19 years). Thirteen subjects (group 1) had mPAp < 25 mmHg (range: 8–24 mmHg), and 14 subjects (group 2) had mPAp ≥ 25 mmHg (range: 25–97 mmHg). All subjects had mean PAWp or LAp < 15 mmHg. We did not exclude any subjects or recordings from the analysis. The demographic and hemodynamic details of the subjects are summarized in Tables 1 and 2. The only statistically significant differences between the two groups were hemodynamic measurements reflecting the presence or absence of PAH. There was no difference in the PAWp, LAp, or cardiac index between the two groups. The two groups did not differ statistically by age, weight, height, body surface area, or body mass index (Tables 1, 2).

Relative power of the frequency band

The postcorrection overall P values showed no significance for this feature, as shown in Table 3.

Sine wave heart sound replicas

The middle panels of Figure 1 show examples of the spectrograms of the original heart sounds (*top*) from a subject with mPAp < 25 mmHg (*left*) and a subject with mPAp ≥ 25 mmHg (*right*) measured at the 2LICS. The bottom panels show the corresponding sine wave replicas. Each sine wave replica has eliminated all extraneous information from the sound file other than the variation of the 4 formants over time.

Feature selection (sinusoid choice) based on energy (Fig. 2)

We investigated which of the 4 sinusoid formants of the heart sounds was most informative. We found that the first sinusoid was the most informative feature for heart sounds collected at the 2LICS.

The energy of the first sinusoid obtained from heart sound recordings at the 2LICS of subjects with mPAp ≥ 25 mmHg was higher than that of subjects with mPAp < 25 mmHg (overall P value = 6.01×10^{-3}). The significance, after statistical postcorrection, was achieved after systematic testing of a range of window lengths of the processed heart sound recording (see "Nonstationarity"). As shown in Table 3, the optimal window length for the energy feature, after optimization, is 5 seconds.

Entropy of the first sinusoid formant derived from the heart sounds (Fig. 3)

The entropy of the first sinusoid formant of the heart sounds recorded at the 2LICS of subjects with mPAp < 25 mmHg was
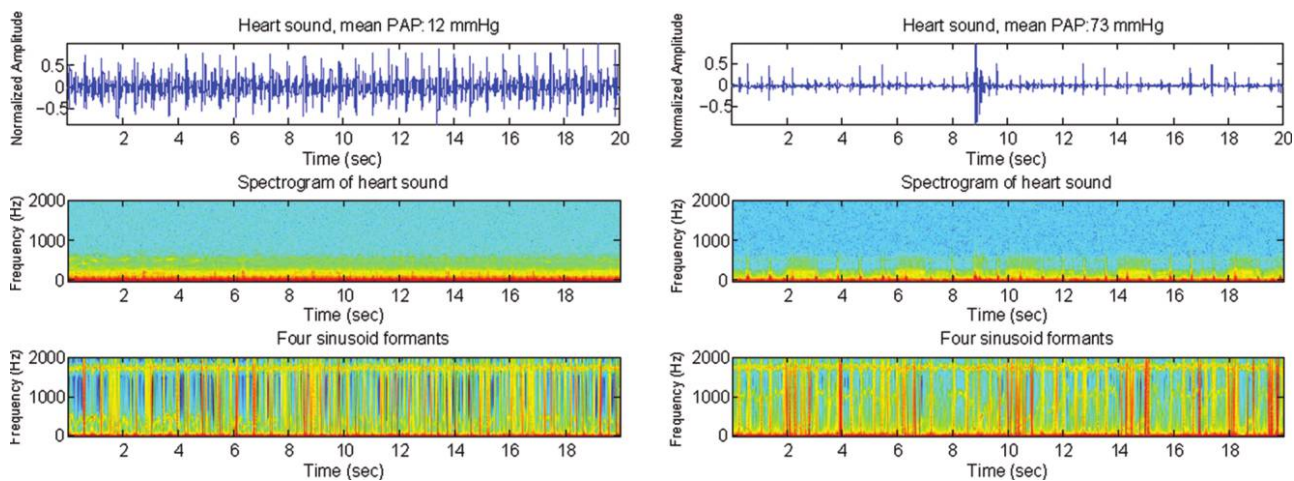

Figure 1. Heart sound recordings (*top*), spectrograms (*middle*), and sine wave replicas that represent the 4 sinusoid formants (*bottom*) of the heart sounds in 2 subjects, one with a mean pulmonary artery pressure (PAp) of 12 mmHg (*left*) and the other with a mean PAp of 73 mmHg (*right*). In the bottom panel, it is hard to distinguish the subject with pulmonary artery hypertension from the subject with normal PAp. This contrasts with the bottom panel of Figure 5, in which the first sinusoid formant demonstrates a clear difference in frequency distribution. *Top*, normalized amplitude (*Y*-axis) plotted against time (*X*-axis); *middle and bottom*, frequency (*Y*-axis) plotted against time (*X*-axis).
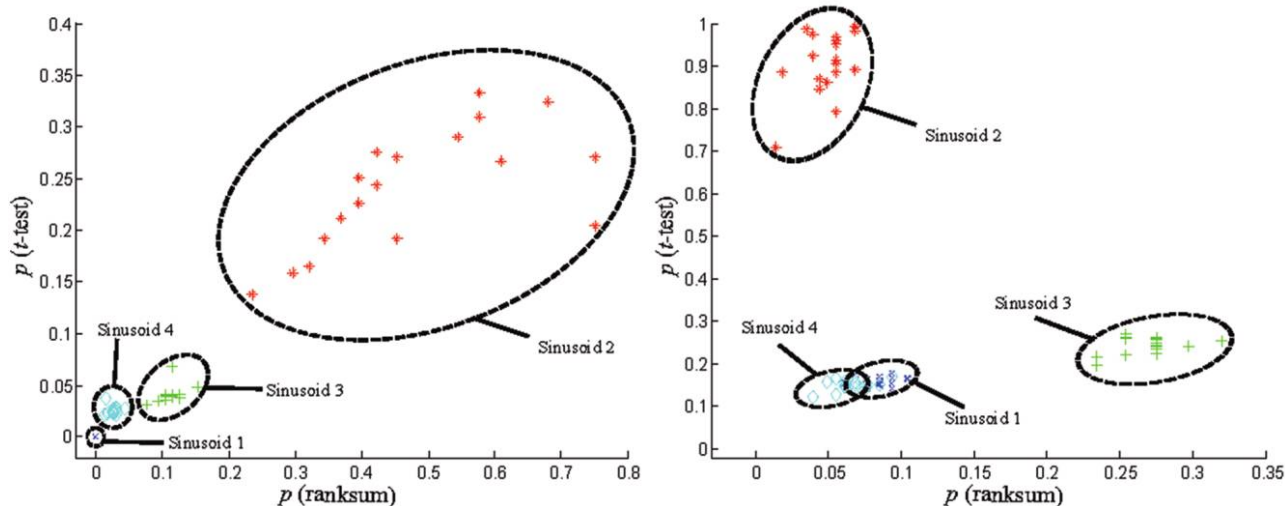
Figure 2. Two-dimensional *P* value representations of the 4 sinusoids that represent formants for heart sounds measured at the second left intercostal space (2LICS; *left*) and the cardiac apex (*right*). The sinusoid formants extracted from heart sound recordings at the 2LICS are separated clearly, whereas at the cardiac apex, they overlap.

significantly higher than that of subjects with mPAp ≥ 25 mmHg (overall *P* value = 3.23 × 10^{-5}). The significance, after statistical postcorrection, was achieved after systematic testing of a range of window lengths of the processed heart sound recording (see "Nonstationarity"). Table 3 shows that the optimal window length for the entropy feature was 2 seconds.

LDA

To ensure that the entropy of the first sinusoid formant of the heart sounds was the most informative feature in PAH, we conducted LDA on the recordings at the 2LICS. Table 4 shows that the entropy of the first sinusoid formant of the heart sounds incurred one false-positive and one false-negative result (Fig. 4). The sensitivity of 93% and specificity of 92% of the entropy of the first sinusoid formant of the heart sounds to detect PAH were superior to both the relative power of the frequency band 21–22 Hz and the energy of the first sinusoid formant (respectively, sensitivity: 71% and 71%, specificity: 69% and 92%).

DISCUSSION

Our main finding was that the entropy of the first sinusoid formant contained within an optimized 2-second window length of the heart sound recordings at the 2LICS was significantly lower in subjects with PAH (mPAp ≥ 25 mmHg), with a sensitivity of 93% and a specificity of 92% (Figs. 3, 4*c*). The reduced entropy of the heart sounds in subjects with PAH suggests the existence of an organized pattern within the heart sounds. A decrease in entropy suggests less chaos and more organization. We have found that the presence of mPAp > 25 mmHg imparts a unique signature to the heart sounds that is reflected by a decrease in entropy within the frequency band 21–22 Hz. Thus, within the frequency range of

21–22 Hz, the heart sounds of subjects with pulmonary hypertension are more organized (less entropy) and demonstrate a different pattern from those of subjects with normal PAp. This pattern could be captured by a noninvasive recording device and used to diagnose PAH. Figure 4 demonstrates that the energy and entropy
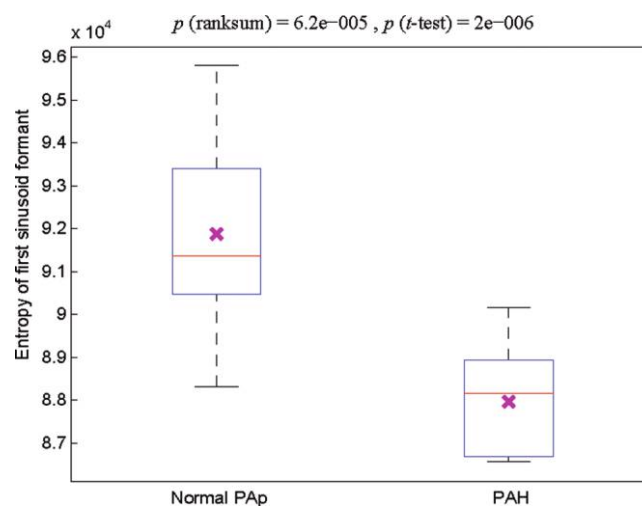

Figure 3. Box plot of the entropy of the first sinusoid formant extracted from the heart sounds recorded at the second left intercostal space. The left-hand box represents the entropy from the heart sounds of children with a mean pulmonary artery pressure (PAp) of 8–24 mmHg (*n* = 13); the right-hand box represents that from the heart sounds of children with a mean PAp of 25–97 mmHg (i.e., with pulmonary artery hypertension [PAH]; *n* = 14). The cross in each box indicates the statistical mean. A two-sample *t* test was performed, and a significant (*P* < 0.05) difference was detected between subjects with a normal PAp and those with PAH. The red line indicates the median. The reported *P* values are uncorrected.

Table 4. Linear discriminant analysis error results, computed through LOO cross validation

Subject	Training error, %	Test label	LDA classification label	Result
1	7.69	PAH	PAH	TP
2	7.69	PAH	PAH	TP
3	7.69	PAH	PAH	TP
4	7.69	PAH	PAH	TP
5	7.69	PAH	PAH	TP
6	7.69	PAH	PAH	TP
7	7.69	PAH	PAH	TP
8	7.69	PAH	PAH	TP
9	7.69	PAH	PAH	TP
10	3.85	PAH	Normal PAp	FN
11	7.69	PAH	PAH	TP
12	7.69	PAH	PAH	TP
13	7.69	PAH	PAH	TP
14	7.69	PAH	PAH	TP
15	7.74	Normal PAp	Normal PAp	TN
16	7.74	Normal PAp	Normal PAp	TN
17	7.74	Normal PAp	Normal PAp	TN
18	7.74	Normal PAp	Normal PAp	TN
19	7.74	Normal PAp	Normal PAp	TN
20	7.74	Normal PAp	Normal PAp	TN
21	7.74	Normal PAp	Normal PAp	TN
22	7.74	Normal PAp	Normal PAp	TN
23	7.74	Normal PAp	Normal PAp	TN
24	7.74	Normal PAp	Normal PAp	TN
25	7.74	Normal PAp	Normal PAp	TN
26	7.74	Normal PAp	Normal PAp	TN
27	7.74	Normal PAp	PAH	FP

Note: FN: false negative; FP: false positive; LOO: leave-one-out; PAH: pulmonary artery hypertension; PAp: pulmonary artery pressure; TN: true negative; TP: true positive.

of the first sinusoid formant were more informative than the relative power of the frequency band 21–22 Hz in distinguishing patients with PAH.[14] Moreover, the entropy of the first sinusoid formant provided better separability between subjects with PAH and those with normal PAp than the energy of the first sinusoid formant. The low entropy suggests that there was an ordered pattern in heart sounds of subjects with PAH that is clearly distinguishable (Fig. 5).

A short recording time of 20 seconds for diagnostic-data acquisition is helpful in real-life clinical settings, particularly in a pediatric clinic when patient cooperation is unpredictable and of limited duration. We used recordings from a digital stethoscope (3M Littmann 3200 electronic stethoscope) and did not exclude any recordings. This suggests that this analysis of the heart sounds in the frequency domain is robust. We speculate that higher-fidelity heart sound sensors would improve the sensitivity and specificity value of the results.

We did not focus our analysis on the detection, timing, or splitting interval between the aortic and pulmonary components of S2. Although these are traditional clinical indicators of PAH, the analysis of the S2, particularly differentiating the aortic and pulmonary components of the S2 and the splitting interval, remains a significant challenge.[6-9,11,12] Therefore, we concentrated on using hidden information within the frequency domain and improved on previous findings by using the relative power of the frequency band 21–22 Hz.[14] We have attempted to characterize the heart sounds of subjects with and without PAH in the same manner as speech patterns and to detect the unique signature of heart sounds in subjects with an increased PAp. This approach is advantageous because it is not necessary to register the timing of heart sound recordings with right heart or pulmonary artery events, which simplifies the approach to noninvasive diagnosis of PAH. It is interesting to note that the recording site that best distinguished patients with PAH was
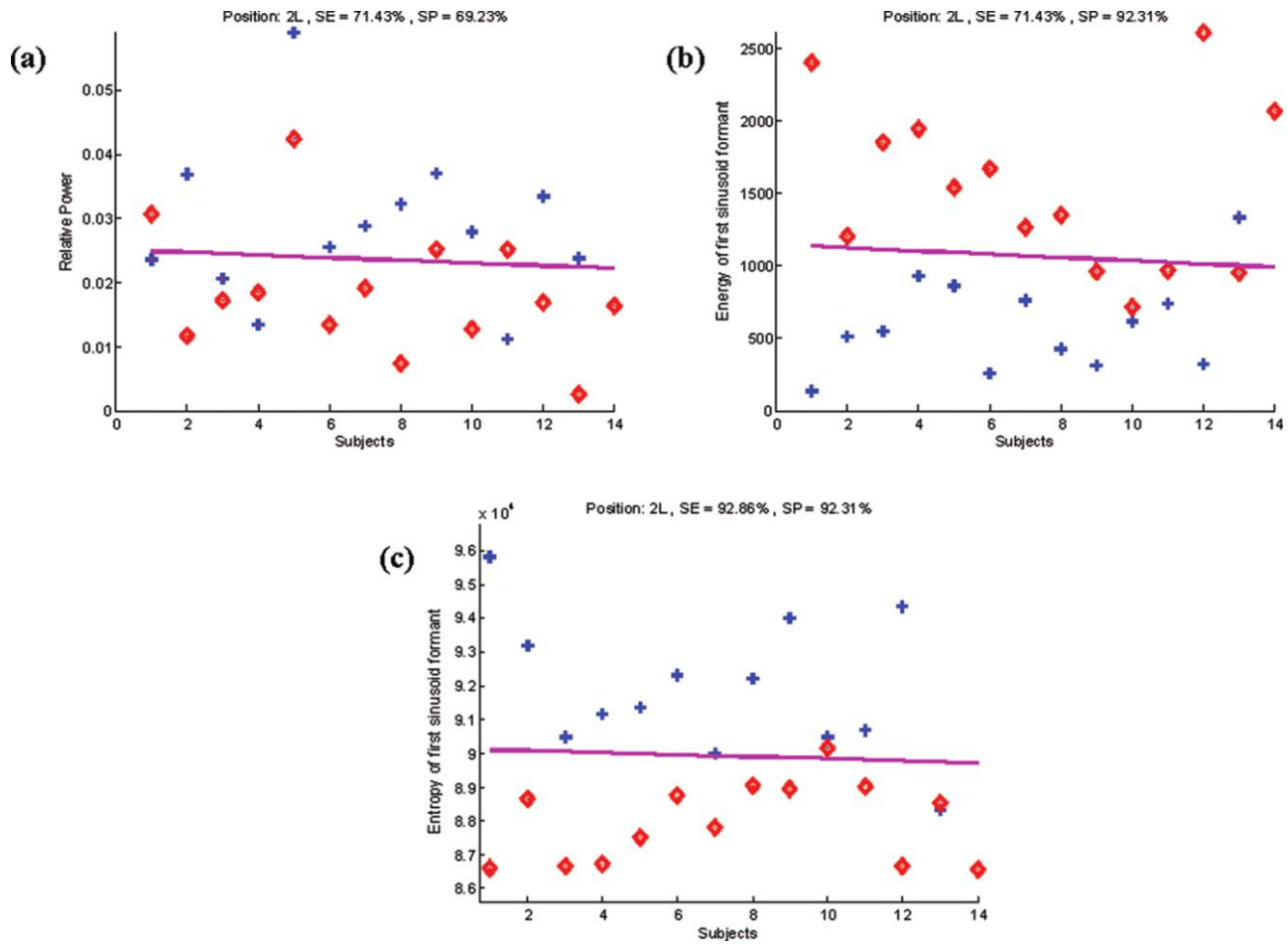
Figure 4. Comparison of three spectral features of the heart sounds (*Y*-axis) for each subject (*X*-axis): relative power of the frequency band 21–22 Hz (*a*), energy of the first sinusoid formant (*b*), and entropy of first sinusoid formant (*c*). Entropy separated subjects with normal pulmonary artery pressure (plus signs) from subjects with pulmonary artery hypertension (diamonds). The pink line signifies the line of separation from linear discriminant analysis. 2L: second left intercostal space; SE: sensitivity; SP: specificity.
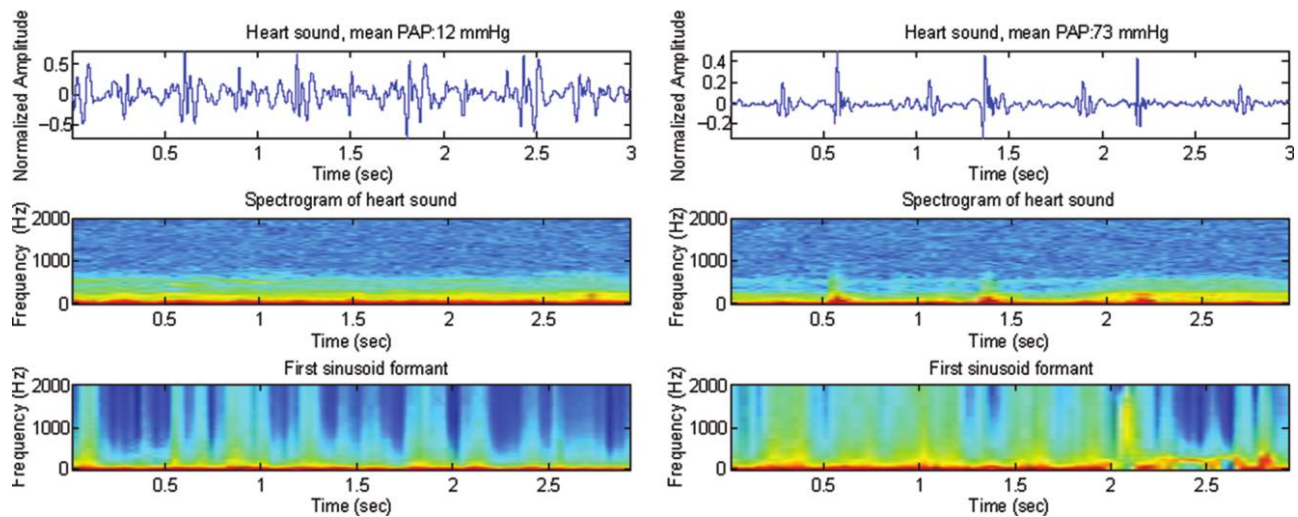

Figure 5. Three-second duration of heart sound recording (*top*), spectrogram (*middle*), and the first sinusoid formant using the sine wave replica (*bottom*) in 2 subjects, one with mPAp = 12 mmHg (*left*) and the other with mPAp = 73 mmHg (*right*). *Top*, normalized amplitude (*Y*-axis) plotted against time (*X*-axis); *middle and bottom*, frequency (*Y*-axis) plotted against time (*X*-axis). In the bottom panel, there is a clear pattern in the frequency distribution representing low entropy in the subject with pulmonary artery hypertension (*right*). mPAp: mean pulmonary artery pressure.

the 2LICS, which is the traditional area for auscultation of pulmonary artery events.

Study limitations

A larger sample size is needed to confirm the findings of this study. We acknowledge that, if this technique were to be applied to a true screening population with a different prevalence of PAH, this might decrease the sensitivity and specificity and that the positive and negative predictivity might be adversely affected.

Prospective recordings with the investigators blinded to the patients' diagnoses and in a population with a lower prevalence of PAH are required in future studies. However, the use of LDA and LOO cross validation to analyze the findings removes investigator bias considerably.

Conclusion

Our data, obtained with a digital stethoscope simultaneously with PAp measurements, showed that the entropy of the first sinusoid formant (within an optimized window length of 2 seconds within a 20-second recording of the heart sounds) at the 2LICS was significantly lower in subjects with PAH, yielding a classification sensitivity of 93% and a specificity of 92%. The reduced entropy of the first sinusoid formant of the heart sounds in subjects with PAH reveals an organized pattern in heart sounds. The analysis of this pattern reveals a unique sound signature produced by the hypertensive pulmonary artery and right ventricle that can be captured and potentially used to diagnose PAH.

REFERENCES

1. Humbert M, Sitbon O, Chaouat A, Bertocchi M, Habib G, Gressin V, Yaïci A, et al. Survival in patients with idiopathic, familial, and anorexigen-associated pulmonary arterial hypertension in the modern management era. Circulation 2010;122(2):156–163.
2. Butrous G, Ghofrani HA, Grimminger F. Pulmonary vascular disease in the developing world. Circulation 2008;118(17):1758–1766.
3. Adatia I, Kothari SS, Feinstein JA. Pulmonary hypertension associated with congenital heart disease: pulmonary vascular disease: the global perspective. Chest 2010;137(6 suppl.):52S–61S.
4. Leatham A. Splitting of heart sounds and a classification of systolic murmurs. Circulation 1957;16(3):417–421; discussion 421–422.
5. Leatham A. Splitting of the first and second heart sounds. Lancet 1954;264(6839):607–614.
6. Longhini C, Baracca E, Brunazzi C, Vaccari M, Longhini L, Barbaresi F. A new noninvasive method for estimation of pulmonary arterial pressure in mitral stenosis. Am J Cardiol 1991;68(4):398–401.
7. Chen D, Pibarot P, Honos G, Durand LG. Estimation of pulmonary artery pressure by spectral analysis of the second heart sound. Am J Cardiol 1996;78(7):785–789.
8. Xu J, Durand LG, Pibarot P. A new, simple, and accurate method for non-invasive estimation of pulmonary arterial pressure. Heart 2002;88(1):76–80.
9. Nigam V, Priemer R. A dynamic method to estimate the time split between the A2 and P2 components of the S2 heart sound. Physiol Meas 2006;27(7):553–567.
10. Elgendi M, Bobhate P, Jain AP, Guo L, Rutledge J, Coe JY, Schuurmans D, Adatia I. Time-domain analysis of heart sound intensity in children with and without pulmonary artery hypertension: a pilot study using a digital stethoscope. Pulm Circ 2014;4(4):685–695.
11. Tranulis C, Durand LG, Senhadji L, Pibarot P. Estimation of pulmonary arterial pressure by a neural network analysis using features based on time-frequency representations of the second heart sound. Med Biol Eng Comput 2002;40(2):205–212.
12. Dennis A, Michaels AD, Arand P, Ventura D. Noninvasive diagnosis of pulmonary hypertension using heart sound analysis. Comput Biol Med 2010;40(9):758–764.
13. Chan W, Woldeyohannes M, Colman R, Arand P, Michaels AD, Parker JD, Granton JT, Mak S. Haemodynamic and structural correlates of the first and second heart sounds in pulmonary arterial hypertension: an acoustic cardiography cohort study. BMJ Open 2013;3:e002660. doi:10.1136/bmjopen-2013-002660.
14. Elgendi M, Bobhate P, Jain S, Guo L, Rutledge J, Coe Y, Zemp R, Schuurmans D, Adatia I. Spectral analysis of the heart sounds in children with and without pulmonary artery hypertension. Int J Cardiol 2014;173(1):92–99.
15. Rich S, Dantzker DR, Ayres NA, Bergofsky EH, Brundage BH, Detre KM, Fishman AP, et al. Primary pulmonary hypertension: a national prospective study. Ann Intern Med 1987;107(2):216–223.
16. Ivy DD, Abman SH, Barst RJ, Berger RM, Bonnet D, Fleming TR, Haworth SG, et al. Pediatric pulmonary hypertension. J Am Coll Cardiol 2013;62(25 suppl.):D117–D126.
17. Hoeper MM, Bogaard HJ, Condliffe R, Frantz R, Khanna D, Kurzyna M, Langleben D, et al. Definitions and diagnosis of pulmonary hypertension. J Am Coll Cardiol 2013;62(25 suppl.):D42–D50.
18. Remez RE, Rubin PE, Pisoni DB, Carrell TD. Speech perception without traditional speech cues. Science 1981;212(4497):947–949.
19. Ellis DP. Sinewave speech analysis/synthesis in Matlab. http://www.ee.columbia.edu/ln/rosa/matlab/sws/. Updated December 4, 2013.
20. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilitá. Pubblicazioni del Regio Istituto Superiore di Scienze Economiche e Commerciali di Firenze 1936;8:3–62.
21. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979;6(2):65–70.