Design and Learnability of Vortex Whistles for Managing Chronic Lung Function via Smartphones

Spencer Kaiser, Ashley Parks, Patrick Leopard, Charlie Albright, Jake Carlson, Mayank Goel[†], Damoun Nassehi[‡], and Eric C. Larson

Southern Methodist University Dallas, TX USA {skaiser,aparks,pleopard,eclarson}@smu.edu Carnegie Mellon University† Pittsburgh, PA USA mayank@cmu.edu

DigiDoc Technologies‡ Egersund, Norway damoun@digidoctech.no

ABSTRACT

Spirometry is the gold standard for managing and diagnosing obstructive lung diseases. Clinical spirometers, however, are expensive and have limited portability. Vortex whistles have shown promise as a potential substitute for clinical spirometers. While vortex whistles are low-cost and are highly portable, only a subset of common spirometry measurements can be measured reliably. Moreover, no research studies have evaluated characteristics of human interaction with vortex whistles, such as maneuver learnability and mental effort. We present a modified 3D-printed vortex whistle design that enables estimation of spirometry measures not previously attainable with traditional vortex whistles. We evaluate the whistle using a pulmonary waveform generator (a commercial standard) and map parameters of the whistle construction to spirometry test endpoints. Through a human subjects trial we evaluate how to personalize whistle parameters for different subjects and assess cognitive workload while using a vortex whistle. We show that, with personalization, vortex whistles are as effective as clinical spirometers for identifying moderate airway obstruction and require similar cognitive load to use.

ACM Classification Keywords

H.5.5. Information Interfaces and Presentation: Sound and Music Computing; H.1.2. User/Machine Systems: Human Information Processing; J.3. Life and Medical Science: Health

Author Keywords

Spirometry; Vortex Whistle; 3D Printing; Cognitive Load;

INTRODUCTION

Chronic disease management typically requires periodic data collection for assessing disease progression, but patient compliance and behavior can pose significant challenges [2, 14]. Technology can play a pivotal role in overcoming these challenges and facilitating data collection for a number of chronic diseases, such as pulmonary ailments like asthma and chronic obstructive pulmonary disease (COPD). For these ailments,

DOI: http://dx.doi.org/10.1145/2971648.2971726



Figure 1. (Top) Our custom vortex spirometry whistle and example user. (Bottom) Pulmonary waveform generator testing environment.

monitoring symptoms and collecting periodic lung function measures is critical for managing treatment and reducing exacerbations [20]. The most accepted form of lung function measurement is known as spirometry—where a subject exhales forcefully through a flow monitoring device [13].

Existing research in ubiquitous health has made strides into low-cost spirometry sensing [15, 16, 9] from mobile phone microphones by processing lip reverberation sounds and vocal resonances. Such systems, however, require machine learning models driven by time-consuming, exhaustive data collection. Moreover, these systems are brittle to changes in the sensing hardware or from sounds different than those trained upon. For healthcare systems, these types of disadvantages are potentially devastating to the intended impact of the technology the accuracy of the health sensing could fluctuate with different mobile phone hardware or from subjects not represented well by the training corpus. Therefore, the technology becomes less reliable and is not adopted as quickly [24, 6].

Finally, spirometry is a difficult, effort dependent maneuver that requires intense concentration and training. Relatively little research has investigated how individuals can learn the spirometry maneuver outside of a clinical setting. To address these research challenges, we investigate the use of a vortex whistle in spirometry, further called Vortex Spirometry. Vortex whistles are special whistles that change their audible frequency in proportion to the flow rate of air passing through

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org. UbiComp '16, September 12 - 16, 2016, Heidelberg, Germany ACM 978-1-4503-4461-6/16/09...\$15.00

them. The idea is not new: research dating back to 1999 [12, 28, 27] has demonstrated that certain spirometry values are viable to be measured using vortex whistles. In 2016, Goel *et al.* [9] investigated the use of 3D-printed vortex whistles for assessing spirometry measures from mobile phones in mildly obstructed and healthy volunteers. While vortex spirometry has been studied for over a decade, whistle design and ease of learning have yet to be systematically investigated.

In this study, we analyze the feasibility of vortex spirometry in ubiquitous healthcare research:

- We present new designs for 3D-printed vortex whistles that employ a tubular side whistle and dynamic flow diverter to more reliably sense flow measurements. Parameters of the design are quantified mathematically for their ability to estimate flow rates during a spirometry maneuver.
- We use commercial standard spirometer testing equipment to test the capabilities of the vortex whistles, enabling the investigation of whistle performance for both healthy and severely ill patients.
- We present robust algorithms for sensing a range of different lung function measures—all of which are capable of running directly on the smartphone (*i.e.*, no need for a cloud infrastructure).
- We evaluate how whistle dimensions can be personalized for individuals based on their expected peak flow rate.
- We evaluate ease of learning spirometry with and without a vortex whistle using metrics of task difficulty and perceived cognitive load.

We conclude that vortex spirometry is effective at detecting moderate and large airway obstruction when personalization and a side whistle are used. While our human subjects study reveals learning to perform a spirometry maneuver with a vortex whistle is not reliable, results also indicate that perceived cognitive load while using a vortex spirometer is comparable to that of a clinical spirometer.

SPIROMETRY FOR UBIQUITOUS HEALTH MANAGEMENT

A vital role of ubiquitous computing is helping use technology to manage chronic diseases. Abowd argues that health management should be at the forefront of applications in ubiquitous and collective computing research:

We are increasingly able to collect clinically meaningful data outside of clinical spaces such as the doctor's office and hospital. Soon, such data will dwarf that received from clinical spaces... [2]

This paper builds upon our previous research, adding novel 3D-printed designs, continued development of algorithms, and extensive evaluation of the technology focused on its role in managing chronic lung diseases.

Chronic Lung Disease Management

Managing chronic lung diseases commonly includes peak flow measurement and symptom monitoring [20]. Daily Symptom Diaries, on the other hand, are an excellent method of improving the management of a wide variety of conditions



Figure 2. Lung obstructions as measured by different spirometry tests. Image manipulated from *cancerpreventiondaily.com*

and diseases. Daily monitoring and recording of symptoms has shown to be beneficial as a method of helping patients manage their symptoms [3] and as a method of assessing the effectiveness of treatment options [10]. Peak flow monitoring is measured as an indicator for large airway obstruction. It is cheap and easy to measure, but has limited predictive power [8, 23].

Next generation management tools will collect spirometry measures that better measure airway inflammation-measures beyond peak flow. To better understand airway inflammation, Figure 2 shows the types of airway sizes. Peak flow measurements do not change with small airway obstruction [17]. This is because the diaphragm and chest wall can move air forcefully through the trachea and bronchi even without the small airways. Not until the larger airways become obstructed does the peak flow begin to measurably reduce. Volume measures, such as the volume exhaled in the first second (FEV_1), change slightly with moderate inflammation of the small airways [17]. However, there must be considerable buildup of small airway obstruction to cause measurable changes in the volume produced. The most accepted measure for small airway obstruction is FEF_{25-75} , the average flow measured between 25% and 75% of the total lung volume. Even so, given that this measure depends on the accurate calculation of the total lung volume, it can be difficult to reliably reproduce [17, 21]. This further highlights the importance of motivating patients to perform maximal-effort spirometry maneuvers [22]. Moving beyond peak flow monitoring is critical to the capture of small airway inflammation and obstruction. Once logged, measures such as FEF_{25-75} can be analyzed with air quality and location to better understand patient triggers. These triggers can remain elusive to individuals that have managed chronic lung disease for decades [26].

Vortex Spirometry vs. Other Methods

Vortex spirometry has a number of advantages over other spirometry sensing techniques. Traditional spirometers are accurate, but costly. Because they are meant for use in the presence of a trained technical coach, they have limited interfaces for giving users feedback and have limited coaching features. For home spirometry, these devices are too costly to deploy except for the most severe cases [5, 14, 4].

Low-cost turbine devices have also been proposed [11]. These devices have typically less accurate volume measurements

and require replacement of the turbine to remain accurate. As such, these devices are not widely used outside of peak flow monitoring. Researchers have investigated the use of turbines with mobile phone gaming [19], although this interface concentrated only on peak flow measurement.

Finally, our previous research [15, 9] used only the mobile phone microphone to infer spirometry measures. That is, users blew at the mobile phone from a short distance and the phone recorded the sound. These systems require a large amount of data collected from a specific microphone configuration across a wide variety of patient demographics. The signal processing and machine learning requires a cloud-connected device to offload the computation. Moreover, the process has been evaluated mostly on healthy volunteers. In follow up studies, we reported a high number of false negatives from the system [9]. As such, the utility of the system for managing diseases is diminished. Participants in the study also had significant errors in reproducing the test, not knowing how far away to hold the phone or how to orient mouth position. Perhaps the most challenging part of the system is its reliance on specific hardware of the device, since amplitude measures from the microphone do not generalize well across a variety of mobile devices. In follow-up work, separate models for each phone type were required, necessitating a time-consuming data collections process for each phone. While amplitude measurement does not generalize well across phone hardware, detecting frequency can generalize well and does not suffer from the non-linear effects of phone position and distance.

Vortex whistles convert airflow into predictable frequency changes. Detecting frequency as a proxy for flow rate is a relatively simple calculation. However, volume measures are still difficult to infer from the flow rate because the vortex whistle must be at a minimum flow rate to generate sound. Moreover, the formation of the vortex inside the whistle takes time to develop. As such, critical measures at the start and end of the test are not directly measurable [9, 7]. To mitigate these problems, we manipulated the design of the whistle to include a tubular whistle and dynamic flow diverter (discussed later).

The vortex whistle, then, has some key advantages over other spirometry sensing methods for use outside the clinic. First, vortex whistles are easy to use with a mouthpiece, obviating the need for users to focus on mouth shape (as with [15, 9]). Second, whistles allow for local processing of the spirometry signal making the design more portable and enabling a UI that can be responsive to the flow rate (*i.e.*, the creation of a game). Third, the additional cost of the whistle is extremely low-less than \$2 USD to 3D print it and even lower to produce at scale. Fourth, frequency detection based sensing allows the method to generalize well across a wide variety of mobile hardware. Finally, the sound generated by the whistle may act as another form of feedback for the user. This is particularly important because users should internalize what a valid spirometry test *feels like*. The dimension of sound may help internalize this process, thus increasing the reliability of the test.

Vortex whistle spirometry is gaining commercial traction. DigiDoc technologies, at the time of writing this paper, is creating a commercial product using a vortex whistle as a peak flow monitor (http://www.digidoctech.no). The beta version of the whistle is made to look like a fish and is geared toward pediatric spirometry. In this paper, we compare performance of our 3D-printed design to the beta version of the DigiDoc fish whistle for measuring peak flow.

PULMONARY WAVEFORM GENERATION

Commercial spirometers have strict requirements for meeting criteria such as acceptable ranges for the measurement of flow, the measurement of scalar values like PEF and FVC, and also the highest acceptable back pressure into the mouthpiece. Moreover, the American Thoracic Society (ATS) has set standards for spirometers to classify poor spirometry tests [18]. A spirometry curve might contain a cough, a hesitation at the beginning of the test, or an early stop (the participant does not exhale for long enough). To ensure that spirometers can accurately meet all criteria, a testing device known as a pulmonary waveform generator (PWG) was created (Figure 1, bottom). A pulmonary waveform generator consists of a large piston that is controlled by a number of actuators and motors-it is about the size of a small refrigerator. Using the PWG, arbitrary flow and volume waveforms can be created, as long as the total volume is less than the total volume of the piston and as long as flow does not exceed the force that the motors can apply to compress the piston.

In this paper, we use a PistonMedical PWG-33 pulmonary waveform generator to generate custom calibration flow profiles for characterizing our vortex spirometry whistles. We also use standardized waveforms from the ATS to evaluate our algorithms for converting the whistle sounds into a continuous inferred flow rate. These curves allow us to simulate flow and volume waveforms from individuals with healthy lung function, asthma, and COPD. Moreover, the PWG uses a motor controller with feedback to generate the curves. This means the controller can adjust the force on the piston to compensate for varying back pressures that may cause different output flow rates than what are specified in the waveform files. These feedback sensors are also logged so that we know the total back pressure exerted from the test and the actual output flow rate that the PWG was able to produce. For low back pressures, the controller can easily compensate to create an identical curve to the input. As the back pressure increases to levels much higher than the specified ATS criteria, the PWG can produce distorted flow rate outputs. When we perform our testing with the PWG, we log the maximum back pressure for each test and whether we can see any visible distortion.

WHISTLE CONSTRUCTION AND OPERATION

A vortex whistle is a specific type of whistle which emits a frequency that is directly affected, and determined by, the flow rate of air passing through it [27]. Typically, vortex whistles consist of a hollowed, cylindrical main chamber, an inlet chamber that is tangential to the main chamber, and a smaller outlet cylinder atop the main chamber. Air enters the main chamber via the inlet chamber, rotates around the circumference of the main chamber, and creates a vortex before it escapes through the outlet (see Figure 3). As air enters the outlet, the rotation speed increases and the air begins to exit the body of the whistle, producing a sound with pitch corresponding to rotation

speed. The pitch of the sound escaping the outlet chamber varies directly with the inlet flow rate, allowing for the direct mapping of frequency to flow rate.

The critical analysis of the vortex whistle is done using the frequency of the sound created from the vortex as it exits the outlet. The frequency can be determined by using measurements of whistle components in the following equation:

$$f = \frac{U}{2\pi RA} \sin\left(\theta\right) \sqrt{\frac{1}{R_f(L + \Delta L)}}$$

where U is the flow rate of air, R is the radius of the vortex in the cylindrical cavity, A is the cross-sectional area of the inlet, θ is related to the angle of the formed vortex, R_f is the average radius of the staying fluid, L is the length of the outlet cylinder, and ΔL is the length of the protruding part of the staying fluid [28]. Therefore, the frequency is linearly proportional to the flow rate, U, and inversely proportional to the cross sectional area of the inlet. Also note that some of the parameters in the above equation are difficult to infer mathematically, necessitating that the linear relationship between flow and frequency be determined empirically [12, 25].

The initial design of our 3D-printed whistle consisted of the previously-mentioned components; a main chamber, an inlet, and an outlet. After testing this original device, it became clear that it would be difficult to obtain several important data-points from the effort. Shortly after beginning an effort, the flow rate is high and the frequencies emitted by the vortex are strong; however, as the vortex begins to form the emitted frequency is weak and undetectable. Similarly, as the flow rate decreases at the end of an effort, the vortex begins to degrade and the emitted frequency becomes undetectable. As a result, the original design was incapable of producing a strong enough frequency at the beginning and end of an effort, preventing the calculation of Forced Vital Capacity (FVC), an inherent issue with standard vortex whistles [7].

To compensate for this pitfall, a side whistle was added to the design. Unlike vortex whistles, aerodynamic whistles generate a strong frequency at low flow rates-an ideal addition to the design to allow for the detection of the low flow rates produced at the beginning and end of an effort. In order to prevent the whistle from interfering with the frequencies generated by the vortex, the length of the whistle's chamber and the diameter of the whistle inlet were modified. Additionally, aerodynamic whistles' frequencies change with pressure [27]. Therefore, when the vortex whistle's pitch is at its maximum, more pressure is created inside the mouthpiece, including the side whistle. This causes the pitch of the side whistle to jump to a higher frequency and avoid interfering with the pitch of the vortex. Through iterative testing, it became apparent that additional modifications must be made to the side whistle parameters to create a pitch which was not irritating to the human ear. After suitable side whistle parameters were identified, the whistle was attached to the side of the device. The inlet was expanded and a splitter was added to divide airflow between the main chamber and the side whistle. Although air would travel through the side whistle, airflow experienced less resis-



Figure 3. Final vortex whistle design including side whistle and side stack. (Blue: side stack; Yellow: mouthpiece; Purple: side whistle; Red: downstream tube; Grey: body). Phases refer to those in Figure 5.

tance traveling through the main chamber than it did traveling through the side whistle. As a result, the frequency emitted from the side whistle was weak and inconsistent.

To solve this problem, we modified the design to force air through the side whistle at low flow rates and through the main chamber at higher flow rates. A chamber containing a weighted ball, referred to as the side stack, was added to top of the device's inlet. The side stack effectively closes the passage into the main chamber until the flow has enough force to lift the ball against gravity and then closes it again when the flow rate decreases enough for the ball to return to its origin. In order to assist airflow in lifting the ball, the angle of side stack chamber was adjusted to a 45-degree slope. After additional design iterations, the chamber was given a conical shape to prevent the imperfect, 3D-printed ball from getting stuck during its ascent or descent mid-effort. With this updated design, the side whistle emits a strong frequency at the beginning and end of each effort while the vortex is able to form and emit a strong frequency immediately preceding and following the time of peak flow. The final design can be seen in Figure 3, along with the annotated airflow for different phases of the tests (beginning:1, middle:2, end:3).

Component Testing

During testing of the whistle components, several important discoveries were made which impacted the final design of the whistle as well as the methodologies of analysis. When testing a whistle comprised of large-sized components, we discovered that when a peak flow of 11 L/s was tested, the back pressure was so significant that the pulmonary waveform generator (discussed in Section 3) would consistently abort the test. A peak flow of 11 L/s is large, but not atypical for an adult male with mid-range height. According to the ATS, a clinical spirometer should be able to measure flows up to 15 L/s [18]. In order to meet these spirometer standards, we needed to make adjustments to the design to accommodate higher flow rates to allow a larger volume of air to travel through the whistle while still forming a stable vortex. To accomplish this, a series of circular vents were added to the side stack, allowing a portion of the air to escape near peak flow. Because these vents were placed above the resting position of the ball in the side stack, air was only able to escape when the flow rate was significant enough to raise the ball above the vents. Furthermore, by limiting





the total area of the vents, a controlled portion of the air is still forced into the main chamber, creating a stable vortex. This design modification allows for a wider range of peak flow values by mitigating back flow while ensuring the creation of a stable vortex.

After completion of the whistle design, an interesting phenomenon was discovered; during an effort, the frequency emitted by the side whistle changes slightly with flow rate. After researching this phenomenon, we discovered that aerodynamic whistles change frequency as the pressure of air entering the whistle changes [27]. Given this finding, it is likely that before the flow rate raises the ball in the side stack and after the ball returns to its origin, the pressure inside the inlet changes, affecting the frequency of the side whistle. Although this finding was unexpected, the additional data point was used as a feature in the machine learning regression.

Modular Component Design

In order for the final whistle design to produce accurate, reliable data for users of different age, height, and sex, the whistle was designed with a modular foundation. With the exception of the whistle base and the side whistle, each of the whistle components have a "Small," "Medium," and "Large" variation that affect the range and strength of the frequency emitted by the vortex as well as the side whistle. Each of the the 3Dprinted components have a tolerance of 0.20 mm, ensuring a firm fit while allowing for easy removal. A variety of whistle components can be seen in Figure 4 and whistle parameters can be seen in Table 1.

	Small	Medium	Large
Down. Tube Length	16 mm	22 mm	28 mm
Inlet Area	45 mm^2	61 mm ²	108 mm ²
Side Stack Area	35π mm ²	45π mm ²	55π mm ²
Ball Weight	0% infill	10% infill	20% infill

Table 1. Table of Whistle Parameters. We also examined the effect of a side stack area of 0 $\mathit{mm}^2.$

Estimating Waveform Criteria from Whistle Dimensions

Using the PWG, we created a series of custom calibration flow waveforms that linearly raised to a peak flow value, sustained peak flow, and then linearly decreased flow to zero. The plateau of these waveforms was increased in steps of 1 L/s from 1 up to 15 L/s (See Figure 5). These calibration files were created so that we could empirically investigate the audio and flow relationship of the vortex whistle as a function of the designed parameters. We set out to determine, for each of the whistle configurations, (A/B) what is the minimum flow rate at which the vortex becomes detectable and stops being detectable, (C/D) at what flow rate does the side whistle become detectable and stops being detectable, (E) what was the maximum back pressure recorded by the PWG for a given flow rate, and (F) at what flow rate does the vortex become turbulent and non-linearly related to frequency? These criteria are enumerated in Figure 5.

For each of the 108 whistle combinations, we recorded the 15 custom waveform files and annotated the resulting 1620 spectrograms for each of the criteria (A-F). This data enabled the creation of polynomial models mapping from the input whistle dimensions to the output criteria. We used regressions that included a bias term, the max input flow rate for a file, each dimension of the whistle, and the interaction terms for each dimension. Therefore, for each whistle we could predict each criterion (A-F) from the whistle dimensions and expected maximum flow rate. We also employed feature selection to reduce the number terms in the polynomial model. Figure 6 shows an example of the actual and predicted maximum back pressure for different whistle configurations and flow rates, with the corresponding regression equation. Two scatter plots are shown, one that uses all parameters for estimation and one that uses only a few parameters.

PERSONALIZING WHISTLE CONSTRUCTION

We investigate if personalizing whistles can provide more accurate spirometry measures for individuals with decreased lung function. In particular, we are interested in knowing if we can minimize the minimum flow rate at which the vortex becomes audible, while also minimizing the highest expected back pressure. We created a personalization algorithm that takes, as input, the expected peak flow rate for an individual and outputs the dimensions of all whistle parameters.

Algorithm: Whistle Dimension Selection

Ideally, a vortex whistle spirometer would (1) begin to have an audible vortex at low flow rates, (2) would remain audible for as long as possible during the forced exhalation, (3) would exert only slight back pressure, and (4) would retain a linear frequency-to-flow relationship up to the maximum peak flow rate for an individual. However, the whistle dimensions for these design criteria contradict one another and, therefore, must be traded off.

Design criteria (3) and (4) can be simplified since minimizing back pressure will also ensure the frequency to flow profile remains linear (especially for the low back pressure rates we strive for in spirometry). Design criteria (1) and (2) are also closely related to one another, although we know that the



Figure 5. The PWG flow profile and spectrogram from whistle audio for a custom flow waveform. This waveform is used to characterize the design parameters of the vortex whistles, which are listed on the right.



Figure 6. Predicted and actual back pressure from empirically determined mapping.

weight of the ball valve has more effect on initial vortex audibility than end of test audibility. The beginning of the test can also many times be inferred from the audibility of the side whistle, which becomes audible at much lower flow rates than the vortex. Because the initial blast of the maneuver occurs in a short period, there are only a few samples to interpret between initial blast and peak flow rate. Therefore, it is easier to interpolate the flow rate, even if the vortex does not become audible until just before peak flow is reached. In contrast, we would like the audibility of the ending vortex to be as low as possible (as long as back pressure is acceptable). We therefore use the following heuristic:

- 1. We create an exhaustive list of the different whistle configurations we investigated and calculate max back pressure and minimum audibility for the vortex beginning and end, denoted as set C_{all}
- 2. Based on the subject's demographic information, we use a peak flow percentile chart to estimate their predicted peak flow, *PEF*
- 3. Using this peak flow rate, we find all whistle configurations with minimum audible vortex end flow rates that are less than one tenth of peak flow, but also maximum peak flow rates above PEF, $C_1 = C_{all} \cap C_{U_{end}} < PEF/10 \cap C_{U_{max}} > PEF$
- 4. We then eliminate all configurations with minimum audible vortex starting flow rates less then one half peak flow $C_2 = C_1 \cap C_{U_{start} < PEF/2}$

5. Finally, we select the whistle configuration with the minimum back pressure, p, from the remaining whistle configurations, $C_{selected} = \arg\min pressure(C_2)$

Using this heuristic, we can be reasonably sure that the vortex sound is adequate for flow estimation during the bulk of the test and minimize the exerted back pressure for the allowable whistle configurations.

Algorithm: Frequency Detection and Regression

For estimating the flow rate, we employ a two step process: (1) frequency and harmonic following and (2) linear multivariate ridge regression. Our frequency following algorithm searches for the expected frequencies within three distinct frequency bands. We use the detected frequencies in each band as features for regressing the actual ground truth flow rate.

An effort from a user results in several continuously generated frequencies. Figure 7 (left) shows a spectrogram of the audio up to 7kHz. Notice that the vortex sound occupies the frequency range from about 300Hz up to 2kHz, the side whistle from about 1.5kHz up to 2kHz, and side whistle harmonic from 5kHz up to 5.5kHz. When the flow rate through the whistle reaches a critical point, the side whistle resonance jumps exclusively to the 5kHz band, as designed. This helps the side whistle and vortex frequencies to not overlap with one another. We start our analysis by computing the spectrogram of the entire audio sequence with sampling rate=44.1kHz, window size=5ms, and overlap=4.3ms (we use rectangular windowing for the highest possible frequency resolution). Each point in the spectrogram is denoted as P(t, f) where t and f are the time and frequency.

We then take the linear sum of each window to estimate the period of the audio with the loudest sound, t_{max} (Figure 7, step 1). We search within $P(t_{max} \pm 20\text{ms}, f)$ for the highest frequency in the range of the vortex whistle, f_{vortex}^{PEF} , resulting in an estimate of where peak flow occurs. We then use $P(t_{vortex}^{PEF}, f_{vortex}^{PEF})$ as a seed point to find the vortex pitch (Figure 7, step 2). We iteratively search forward in time over a range of 20Hz for the next largest magnitude, $P(t_{seed} + 1, f_{seed} \pm 20\text{Hz})$, update the seed point, and repeat. We stop search when the magnitude of the resonance is sufficiently small, $P(t_{seed}, f_{seed}) < 0.2P(t_{vortex}^{PEF}, f_{vortex}^{PEF})$. We then repeat the process from the original seed point, but this



Figure 7. Overview of the frequency following algorithm and features extracted for regression.

time working backwards in time, $P(t_{seed} - 1, f_{seed} \pm 20 \text{Hz})$ as shown in Figure 7, step 2.

The times where the vortex frequency stops tracking, t_{vortex}^{end} and t_{vortex}^{start} are then used as new seed points in the frequency band of the side whistle (Figure 7, step 3). The frequency tracking is performed identically as before, but now using two seed points to track the discontinuous resonances. Finally, we track the harmonic of the side whistle using a seed point of t_{max} identically to the method used for following the vortex frequency (Figure 7, step 4). These three tracked resonances are then saved as features to predict the ground truth flow rate generated via the PWG (Figure 7, step 5).

To obtain ground truth the linear regression, we use the input flow rate file from the pulmonary waveform generator. We resample the sampling rate of the flow waveform to be at the same as the input features (the input features have an effective sampling rate of about 300Hz and the waveform file is 100Hz). Finally, we align the curves based upon the time the vortex reaches maximum frequency and when the peak flow occurs in the waveform.

In addition to using the personalized, 3D-printed whistles, we also used an injection molded vortex whistle designed to measure peak flow. This whistle is an early prototype of a commercial product to be released by DigiDoc Technologies. This whistle does not use a ball or side whistle, so we manipulated our frequency detection algorithm to only look for the fundamental vortex frequency. The regression is based solely from this single tracked resonance.

Results: Back Pressure Comparison

To test the whistles we connect different whistle configurations to the pulmonary waveform generator and record audio of standardized curves from the ATS. This results in 50 curves (the same set of curves used to test commercial spirometers). For each of the 50 curves, we collect data using a personalized whistle configuration and a "one-size-fits-all" configuration. The "one-sized" whistle was chosen to be mid-sized—to trade-off sensitivity and back pressure. During the testing, maximum back pressure and any visible distortion from the waveform generator was logged by the experimenter.

Figure 8 shows distribution plots of the maximum measured back pressure during the testing. When using a "one-sized" whistle configuration it is apparent that back pressures are significantly higher than those of a personalized configuration. Personalized whistles resulted in curves with back pressures largely below 2kPA. While this is comparatively lower than the "one-sized" whistle, it is higher than commercial spirometers, which have back pressures less than 1kPA. Also shown are the number of tests for which the waveforms were not visibly distorted. For personalized whistles, no visible distortion was observed. The "one-sized" whistle only passed 41 of the 50 tests. This highlights the importance of adjusting whistle dimensions to prevent curve distortion.

Results: Lung Function Measures

To test the personalization protocol, we create separate regression models for each of the unique whistle configurations used. Our personalization algorithm selected 5 different whistle combinations across a range of peak flows from 0.5 to 14.6 L/s. For each whistle configuration, we cross validate the prediction using leave-one-curve-out cross validation. That is, we hold out each unique curve from the pulmonary waveform generator and train the regression on the remaining curves. The cross validation results in a predicted flow curve for each test. Once we have a predicted flow rate, we can calculate the volume curve by integrating the values. These two curves allow the calculation of any exhaled spirometry measure. We



Figure 8. Distribution plot of maximum back pressure for personalized and "one-size-fits-all" whistle configurations.

report the results for PEF, FEV₁, FVC, and FEV₁%. To investigate the effect of personalization and the effect of the side whistle, we train a regression model using the "one-sized" whistle (A) and we train a regression model without using frequency information from the side whistle (B). These are compared to the personalized whistle configuration with a side whistle (C). Finally, we compare our results to the injection molded DigiDoc whistle (Digi). Because the DigiDoc whistle is specifically designed for measuring PEF, we only compare the results for PEF.

Figure 9 shows boxplots of percentage difference for each of the measures, $\frac{\hat{y}-y}{y}$ where y is the ground truth measure and \hat{y} is the predicted measure. The FEV₁/FVC measure is not normalized, but is instead the raw difference (as it is already a percentage). Each whistle configuration is grouped according to A, B, C, and Digi, as described. Also shown are the number of tests in each configuration that produced curves without visual distortion (as measured by the pulmonary waveform generator) and produced an audible vortex sound. Without personalization (A) only 34 of the 50 curves produced an audible vortex and distortion free curves. With personalization (group B and C) 46 curves were valid. The 4 curves that failed had peak flow rates less than 1 L/s and our personalization algorithm failed to find a configuration supporting such low flow rates. For the Digi group, 26 of the 50 curves were valid. We also conducted an F-test of the residual variances between each of the groups (the Jarque-Bera test revealed all percentage residuals were sufficiently Gaussian to perform an F-test). For PEF and FEV₁/FVC, no group had significantly different variances (p>0.05). For FEV₁ and FVC, group C had statistically smaller variance than all other groups (p < 0.01). Also for FVC, group A had significantly smaller variance than group B. The impact of personalization was most apparent for FEV₁ with group C having an interquartile range (IQR) nearly half the size of other groups. Moreover, the effect of the side whistle is most apparent for estimating FVC, with groups A and C having significantly smaller IQR than group B. These results support a conclusion that personalization and the side whistle are warranted for measuring volumetric spirometry measures. Even so, a "one-size" whistle could still have utility in diagnostic screenings, where it is not feasible to carry many different sized pieces.



Figure 9. The percentage difference between ground truth and the predicted spirometry measures.

The injected molded DigiDoc whistle had superior estimation of PEF (p<0.05). However, only 26 curves of the 50 produced an audible vortex and distortion free curves. The personalized whistles, on the other hand, could adjust their flow profiles to account for uncharacteristically high or low flows. We also investigated the predictions for FEF₂₅₋₇₅, but the results are not reliable. The percentage difference was many times above 10% from ideal. This is because the measurement relies on accurate peak flow estimation and accurate FVC estimation in the same curve. The current whistle designs are not well suited for measuring FEF₂₅₋₇₅. This also means that more work must be carried out to claim that vortex spirometry is a reliable management tool for tracking small airway obstruction. Based upon these results, vortex spirometry may only be indicated for tracking large and moderate obstruction via PEF and FEV₁.

HUMAN SUBJECTS TRIAL

Several studies have been conducted which focus on the evaluation of vortex whistles. However, the majority have focused on the evaluation of channel mediums [16] or the analysis of their reliability and accuracy [28]. Up until this study, however, no study has evaluated human interaction with vortex whistles. There are many unknowns about the human perspective on vortex whistles, such as how they are perceived and their ease of use in terms of both learning how to use the whistle and how to continue to use it over time.

Experiment Design

The core of the IRB approved human subjects trial centered around evaluating the subjective cognitive load experienced by participants as they performed a full spirometry effort. Cognitive load is a good indication of task difficulty for tasks requiring physical effort, mental effort, and coordination. As such, in our studies we want to quantify perceived difficulty in learning to perform spirometry with and without a vortex whistle—cognitive load can therefore be a helpful indicator. Furthermore, we wanted to gather information about how coaching affected overall cognitive load (*i.e.*, is it more difficult to learn spirometry from a whistle than from a clinical spirometer?). Participants were divided into three groups:

- **Control Group**: an initial spirometry effort is performed using a clinical spirometer with experimenter coaching provided before, during, and after each effort.
- Whistle with Experimenter group: an initial spirometry effort is performed using our 3D-printed vortex whistle with

experimenter coaching provided before, during, and after each effort.

• Whistle with Digital Coaching group: an initial spirometry effort is performed using our 3D-printed vortex whistle. Coaching is provided exclusively via a smartphone application. No human experimenter coaching is provided.

The trial utilized several types of spirometers; a clinical spirometer, our 3D-printed vortex whistle, and the injection-molded DigiDoc Asthma Whistle. While the clinical spirometer performed analysis on the device itself, a smartphone application was developed to handle audio recording and pre-processing for use with the both vortex whistles.

Initial Procedures

Figure 10 shows the timeline of steps for each group. After reading and signing the study consent form, participants were given a brief demographic survey consisting of information such as age, sex, height, weight, whether they had been diagnosed with asthma or COPD, and whether they smoked. Some of this data were used with the clinical spirometer for estimation of expected performance values. In addition, the survey included several questions pertaining to exclusion criteria, such as whether the participant had used a rescue inhaler recently.

Initial Spirometry Test

Participants in the **control group** were given a brief introduction to spirometry and were instructed on the usage of the NDD EasyOne[®] Plus clinical spirometer. Throughout the duration of the test, the experimenter provided feedback, encouraging the participant to focus on the initial blast of the effort and to continue exhaling until the spirometer provided feedback to end the effort. If participants were physically able, this process continued until 3 reproducible efforts had been performed.

Participants in the **Whistle with Experimenter group** were given a 3D-printed vortex whistle and an experimenter provided an introduction to spirometry and how to perform an effort. The experimenter would then introduce the participant to the 3D-printed whistle as well as the smartphone app. During efforts, the experimenter was unable to provide audible coaching to avoid interfering with audio recordings. However, coaching gestures, such as hand gestures and mouthing words, were used to encourage the participant to exhale fully and continue for the full duration of an effort.

Participants in the **Whistle with Digital Coaching group** were also given a 3D-printed vortex whistle, however, no introduction to spirometry or coaching was performed by the experimenter. Instead, participants in this group were asked to utilize a digital coaching system embedded within the smartphone app. Digital coaching consists of three primary sections: background info, training, and testing. The background info portion discusses the importance of measuring lung function and the basics of spirometry. Training goes into detail regarding how to perform a spirometry effort and test, common issues to try and avoid, and ends with a video demonstrating how to use the 3D-printed vortex whistle with the smartphone app. Finally, the participant takes a brief quiz covering all the



Figure 10. Left: Overview of the experimental timeline for each group. Right: question with multiple-choice-answers and response feedback used in the Digital Coaching group.

material presented to them. After completing the coaching, the participant is asked to complete a spirometry test with the whistle without asking the experimenter spirometry-related questions.

Subjective Cognitive Load Evaluation

At the end of the initial spirometry test, participants of all groups were asked to complete the NASA Task Load Index (TLX) regarding their use of the initial device (i.e., a spirometer or whistle depending on the group). The TLX measures workload based on six subscales: mental demands, physical demands, temporal demands, performance, effort, and frustration [1]. Data analysis for the TLX is performed by identifying weight values and then using those values in conjunction with ratings for each of the subscales to calculate an overall subjective cognitive workload. The values from this analysis are used to determine the perceived cognitive load of learning spirometry in each group. Subjects are asked to focus on the task of learning and then performing the spirometry maneuver while completing the NASA TLX.

Additional Testing

After completion of the NASA TLX, participants in the control group were given a 3D-printed whistle and participants of the experimental groups were given the clinical spirometer. Using the provided device, the participant then performed a full spirometry effort with coaching from the experimenter.

Following the completion of the second test, all participants are asked to perform a final spirometry test using the DigiDoc Asthma Whistle. During this test, participants have typically developed a firm understanding of how to perform an effort, but are still coached by the experimenter through the test.

The clinical spirometer provides automatic grading of test quality (A through F). Three reproducible maneuvers are required to attain a grade of "C" and, depending on other criteria such as reaching peak flow quickly and sustaining the maneuver for six seconds, participants can achieve higher grades. If participants were unable to produce a session quality of C or better on the clinical spirometer for reasons other than physical limitation, their data were excluded from the study. This occurred for two participants that would not follow experimenter feedback and appeared unmotivated to participate.

Exit Survey

Upon completion of the final spirometry test, participants were asked to complete an exit survey. The survey consisted of several questions pertaining to their experience performing spirometry tests, whether they would prefer to use a clinical spirometer or a vortex whistle if asked to perform a test daily, and which of the utilized vortex whistles was preferred. Several analyses of this data can be seen in the following section.

Results

We enrolled 21 participants, two of which were excluded for failing to perform a reproducible spirometry maneuver. Another participant met exclusion criteria because of recent use of a fast acting inhaler. Of the remaining 18 participants, ten were male and eight were female. Each of the three study groups consisted of six participants.

One of the focal points of the study was evaluating cognitive load. Overall cognitive load data from the NASA TLX, which is represented by a value between 0 and 100, was calculated for different groupings which can be seen in the swarm and box plots in Figure 11. Also shown are the level of frustration users reported and the perceived performance at completing the task. The results indicate that subjective cognitive load was relatively consistent across all study groups and that the use of a vortex whistle marginally increased cognitive load compared to the clinical spirometer. Although sample size is small increasing the likelihood of erroneously concluding there is no difference between groups, we can qualitatively see that the cognitive load from using a vortex whistle is not unduly high compared to a clinical spirometer.

Perceived performance on the spirometer was much lower than other groups, likely because of the feedback given to the users by the clinical spirometer. Perceived performance was highest in the Whistle and Digital Coaching group, most likely because participants did not receive feedback from the experimenter when they performed an invalid effort (*i.e.*, they were unaware of poor performance). These observations also explain the difference in frustration levels—the feedback from the spirometer, while helpful in gauging performance, may be a source of frustration.

In addition to cognitive load, another focus of the human subjects trial was to evaluate the ability of participants to learn how to perform a reproducible spirometry test (*i.e.*, a grade of C or better). For subjects that learned the maneuver on the vortex whistle (*i.e.*, subjects not in the control group), we asked them to compare their initial performance on the whistle after they had been coached to perform the maneuver with the clinical spirometer. All subjects reported they internalized the spirometry maneuver differently after getting feedback from the clinical spirometer. Given the chance to repeat the tests with the whistle, participants stated that they would have given a more maximal effort. Many participants mentioned that the audible feedback from the clinical spirometer was helpful for understanding when the effort had ended—they suggested emulating this in our smartphone application. From



Figure 11. Swarm and box plot of NASA TLX participant responses for perceived performance, frustration, and overall weighted cognitive load.

this result, we can conclude that subjects are more likely to internalize a valid effort if they learn on a clinical spirometer. More research is required to understand if vortex whistles can be used reliably at home after learning how to perform a valid effort. Learning spirometry through a digital coach is also an open research topic as our digital coaching protocol appeared to be ineffective.

Roughly 65% of participants reported that they preferred one of the vortex whistles over the clinical spirometer. When asked which of the two whistles participants preferred, 72% preferred the DigiDoc Asthma Whistle over the 3D-printed whistle, suggesting the user experience of our whistle can be improved. When asked for additional feedback about this choice, participants mentioned aesthetics like the rubber grip and "solid feel" of the DigiDoc whistle.

CONCLUSION

We presented a novel 3D-printed vortex whistle design and evaluated the design using a commercial pulmonary waveform generator. Moreover, we conducted human subjects test to show the difficulty of the whistle compared to traditional spirometry. We conclude that vortex whistles are low-cost, portable alternatives to clinical spirometers for managing moderate airway obstruction. Our whistle design underwent multiple iterations and evolved from a basic vortex whistle design to one that can accommodate a wide range of flow rates. Additionally, the final whistle design allows for the post-hoc identification of the beginning and end of an effort; data that can be used to identify moderate airway obstruction.

Our human subjects trial suggests that users' cognitive load when using our whistle is comparable to a clinical spirometer. Moreover, data collected from the exit survey revealed that nearly two-thirds of all participants preferred a vortex whistle.

In the future, a longitudinal study evaluating the use of a vortex whistle and associated smartphone app may open additional doors for studies investigating the impact of performing daily spirometry tests on the management and treatment of asthma. This study could focus on the identification of triggers and, when combined with the use of daily symptom diaries, may provide physicians with valuable, long-term data collected on a daily basis about the lives of patients living with asthma; information which is currently unattainable.

REFERENCES

- 2006. NASA TLX: Task Load Index. (2006). http://humansystems.arc.nasa.gov/groups/tlx/
- Gregory Abowd. 2016. Beyond Weiser: From Ubiquitous to Collective Computing. *IEEE Computer Magazine* (2016), 17–23.
- 3. Christina Baggott, Faith Gibson, Beatriz Coll, Richard Kletter, Paul Zeltzer, and Christine Miaskowski. 2012. Initial evaluation of an electronic symptom diary for adolescents with cancer. *JMIR research protocols* 1, 2 (2012).
- 4. Yvonne Bastian-Lee, Richard Chavasse, Hazel Richter, and Paul Seddon. 2002. Assessment of a low-cost home monitoring spirometer for children. *Pediatric pulmonology* 33, 5 (2002), 388–394.
- Alwin F. Brouwer, Ruurd Jan Roorda, and Paul L. Brand. 2006. Home spirometry and asthma severity in children. *European Respiratory Journal* 28, 6 (2006), 1131–1137.
- 6. P. Chowienczyk, D. Parkin, C. Lawson, and G. Cochrane. 1994. Do asthmatic patients correctly record home spirometry measurements? *BMJ* 309, 6969 (1994), 1618.
- Yan-bin Di, Christof Gerhardy, and Werner Karl Schomburg. 2013. Vortex whistles employed as remote micro flow sensors. *Flow Measurement and Instrumentation* 34 (2013), 11–18.
- Jonathan M Feldman, Haley Kutner, Lynne Matte, Michelle Lupkin, Dara Steinberg, Kimberly Sidora-Arcoleo, Denise Serebrisky, and Karen Warman. 2012. Prediction of peak flow values followed by feedback improves perception of lung function and adherence to inhaled corticosteroids in children with asthma. *Thorax* 67, 12 (2012), 1040–1045.
- Mayank Goel, Elliot Saba, Maia Stiber, Eric Whitmire, Josh Fromm, Eric C. Larson, Gaetano Borriello, and Shwetak N. Patel. 2016. SpiroCall: Measuring Lung Function over a Phone Call. In *International Conference* on Human Factors in Computing Systems, CHI'16. ACM.
- C. Goss, T. Edwards, B. Ramsey, M. Aitken, and D. Patrick. 2009. Patient-reported respiratory symptoms in cystic fibrosis. *Journal of Cystic Fibrosis* 8, 4 (2009), 245–252.
- 11. Siddharth Gupta, Peter Chang, Nonso Anyigbo, and Ashutosh Sabharwal. 2011. mobileSpiro: accurate mobile spirometry for self-management of asthma. In *Proceedings of the First ACM Workshop on Mobile Systems, Applications, and Services for Healthcare.* ACM, 1.
- 12. Kajiro Watanabe Hiroshi Sato, Masayuki Ohara and Hideaki Sato. 1999. Application of the Vortex Whistle to the Spirometer. *Transactions of the Society of Instrument and Control Engineers* 35, 7 (1999), 840–845.
- 13. Romain Kessler, Elisabeth Stahl, Claus Vogelmeier, John Haughney, Elyse Trudeau, Claes-Goran Lofdahl, and

Martyn R Partridge. 2006. Patient understanding, detection, and experience of COPD exacerbations: an observational, interview-based study. *CHEST Journal* 130, 1 (2006), 133–142.

- 14. Santosh Kumar, Wendy J Nilsen, Amy Abernethy, Audie Atienza, Kevin Patrick, Misha Pavel, William T Riley, Albert Shar, Bonnie Spring, Donna Spruijt-Metz, and others. 2013. Mobile health technology evaluation: the mHealth evidence workshop. *American journal of preventive medicine* 45, 2 (2013), 228–236.
- Eric C Larson, Mayank Goel, Gaetano Boriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N Patel. 2012. SpiroSmart: using a microphone to measure lung function on a mobile phone. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 280–289.
- Eric C Larson, Mayank Goel, Morgan Redfield, Gaetano Boriello, Margaret Rosenfeld, and Shwetak N Patel. 2013. Tracking lung function on any phone. In *Proceedings of the 3rd ACM Symposium on Computing for Development*. ACM, 29.
- 17. William McNulty and Omar S Usmani. 2014. Techniques of assessing small airways dysfunction. *European Clinical Respiratory Journal* 1 (2014).
- Martin Raymond Miller, JATS Hankinson, V Brusasco, F Burgos, R Casaburi, A Coates, R Crapo, P Enright, CPM Van Der Grinten, P Gustafsson, and others. 2005. Standardisation of spirometry. *European respiratory journal* 26, 2 (2005), 319–338.
- Shawn Nikkila, Gaurav Patel, Hari Sundaram, Aisling Kelliher, and Ashutosh Sabharwal. 2012. Wind runners: designing a game to encourage medical adherence for children with asthma. In CHI'12 Extended Abstracts on Human Factors in Computing Systems. ACM, 2411–2416.
- 20. Klaus F Rabe, Suzanne Hurd, Antonio Anzueto, Peter J Barnes, Sonia A Buist, Peter Calverley, Yoshinosuke Fukuchi, Christine Jenkins, Roberto Rodriguez-Roisin, Chris Van Weel, and others. 2007. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *American journal of respiratory and critical care medicine* 176, 6 (2007), 532–555.
- 21. Devika R Rao, Jonathan M Gaffin, Sachin N Baxi, William J Sheehan, Elaine B Hoffman, and Wanda Phipatanakul. 2012. The utility of forced expiratory flow between 25% and 75% of vital capacity in predicting childhood asthma morbidity and severity. *Journal of Asthma* 49, 6 (2012), 586–592.
- H. Reddel, S. Ware, C. Salome, C. Jenkins, and A. Woolcock. 1998. Pitfalls in processing home electronic spirometric data in asthma. *European Respiratory Journal* 12, 4 (1998), 853–858.
- 23. Attilio D Renzetti Jr. 1979. Standardization of spirometry. *American Review of Respiratory Disease* 119, 5 (1979), 693–694.

- 24. Marti L Riemer-Reiss and Robbyn R Wacker. 2000. Factors associated with assistive technology discontinuance among individuals with disabilities. *Journal of Rehabilitation* 66, 3 (2000), 44.
- 25. Hiroshi Sato and Kajiro Watanabe. 2000. Experimental study on the use of a vortex whistle as a flowmeter. *Instrumentation and Measurement, IEEE Transactions on* 49, 1 (2000), 200–205.
- A. Ten Brinke, P. Sterk, A. Masclee, P. Spinhoven, JT Schmidt, A. Zwinderman, K. Rabe, and E. Bel. 2005. Risk factors of frequent exacerbations in difficult-to-treat asthma. *European Respiratory Journal* 26, 5 (2005), 812–818.
- 27. Bernard Vonnegut. 1954. A vortex whistle. *The Journal* of the Acoustical Society of America 26, 1 (1954), 18–20.
- 28. Kajiro Watanabe and Hiroshi Sato. 1994. Vortex whistle as a flow meter. In *Instrumentation and Measurement Technology Conference*, 1994. IMTC/94. Conference *Proceedings. 10th Anniversary. Advanced Technologies in I and M., 1994 IEEE*. IEEE, 1225–1228.