Ambulatory monitoring of respiratory effort using a clothing-adhered biosensor

Mark Holt, Ben Yule, Dylan Jackson, Mary Zhu, Neema Moraveji

Spire, Inc.
San Francisco, CA, USA
mark, ben, dylan, mary, neema@spire.io

Abstract— Accurate ambulatory sensing of patterns of respiratory effort has long eluded biomedical researchers. Though the data would inform clinical decision-making, the lack of sensors suitable for such data collection stymies clinical impact. This study describes an approach to sensing respiratory effort via thoracic and/or abdominal excursion in a form that affords longitudinal and continuous adherence. This approach leverages force-based sensors embedded in a clothing-adhered form factor to reduce user inconveniences. The primary benefit of this approach is user acceptance: it can monitor data longitudinally while addressing impactful user inconvenience issues with existing devices. Compared to a ground truth monitor of respiratory effort, and across both cognitive and physical tasks, the present approach resulted in a relative median error of 6.8% and mean absolute error of 1.8 breaths per min (SD=0.14). Sensor location affected performance, with chest-worn sensors outperforming waist-worn sensors. A more granular analysis of temporal markers of the respiratory cycle showed high agreement with ground truth; end-of-expiration temporal markers exhibited the least precision. The results indicate that this approach can be used to monitor respiratory effort accurately and theoretically with high adherence.

Keywords—respiration, sensor, wearable, clothing, monitor

I. INTRODUCTION

In physiological monitoring research, the question of data capture has evolved from how to capture accurate data to how to do so in practice, in a way so as to yield sufficiently high adherence over time to reveal insights that have yet eluded clinical decision-making and patient utility. In the present case, we explore respiratory patterns, a fundamental autonomic which can indicate cognitive, emotional, or physiological disorder [9], [26], [29], [20] [28]. Respiratory distress is a primary factor of hospital (re)admission in congestive heart failure (CHF) and chronic obstructive pulmonary disease (COPD) [2] and respiratory feedback can be used as a preventative intervention in asthma [17], anxiety [14], hypertension [22], and more.

Despite its fundamental clinical import, respiratory patterns have been notoriously problematic to capture longitudinally outside controlled or sleep settings. This is attributed to noise arising from physical movements, physical discomfort of existing sensor form factors, behavioral changes required in remembering to wear and recharge sensors, and the wide variation in naturally-occurring breathing patterns (e.g., labored, speaking, laughing, breath-holding, etc.) [28]. Though respiratory rate (RR) is most commonly reported as the primary respiratory feature of interest, the respiratory effort waveform measured directly from respiratory mechanics (i.e., rather than inferred from other biological rhythms) includes other important clinical features such as respiratory stability, inspiration-to-expiration ratio, relative minute ventilation, expiratory duration, and events such as coughs and apneic events. The lack of continuous sensing and reporting of these features impedes clinical decision-making and patient outcomes. Thus, we are interested in exploring methods of delivering this data in a form factor that affords longitudinal patient adherence.

II. RELATED WORK

This section explores contact and non-contact methods of directly assessing respiratory effort (i.e., not indirectly through other physiological data like blood oxygen saturation or cardiac rhythm) for the goal of longitudinal data collection on the scale of months. Placing a thermistor by the nose, for example, was here deemed too obtrusive for longitudinal use. Generally speaking, all monitoring methods balance trade-offs between discomfort, accuracy, and hygiene [1].

Perhaps the most common approach of direct monitoring of respiratory effort is to affix respiratory inductance plethysmography bands either directly to the body or into a skin-tight, torso-worn harness or shirt [8], [13][25]. This involves using one, or ideally two, band(s) around the torso and/or abdomen to gauge respiratory effort as the torso expands and contracts with inspiration and expiration. The bands are constructed from an extendable/deformable conducting material, such as a very fine wire or thin foil such that conductivity is maintained while stretching.

Another contact-based approach is clinician-positioned single and multiple electrodes on the chest wall to measure transthoracic impedance alterations caused by respiratory effort [26]. These approaches have also employed triaxial accelerometers as a means of noise-cancelation [5]. These will require regular maintenance and repositioning and are most appropriate for short-term monitoring.

Acoustic data can be collected at the neck [6] or near the nose [27] though these methods are ostensibly not suited for continuous use due to physical discomfort and social awkwardness.

Non-contact approaches can be employed for longitudinal collection but suffer from being context-sensitive (i.e., the patient must be still and within the range of the sensor). These are best suited for monitoring respiration during sleep. Doppler and thermal [11] techniques have been used to measure...
respiration. Optical techniques (with and without projected infrared lights onto the subject) have also been used to deduce respiratory trends from the movement of the chest wall during respiration [4][19].

Though there have been many approaches to directly monitoring respiratory effort, continuous and longitudinal sensing has not yet been adequately addressed.

III. A NOVEL FORM FACTOR FOR MONITORING RESPIRATORY EFFORT

There are numerous obstacles to high adherence for wearable biosensors such as remembering to charge a device, finding the device each morning, physical discomfort on one’s wrist or chest, needing to reposition it with precision, skin irritation, self-consciousness about a visible health monitor, and more [16]. Until these points of inconvenience are eliminated, continuous and longitudinal patient monitoring will be stymied.

The Health Tag (Spire, Inc., San Francisco, CA) is the first known instantiation of a device category with the same name, health tag: any clothing-attached device with on-board sensors that attempts to attain high adherence and proximity with the wearer’s vitals by being coupled with garments rather than being a stand-alone device. They can be attached in many ways and to various types of garments. Health Tags are sealed and temperature resistant to survive washer and dryer cycles while attached to a garment.

From the patient’s perspective, the Health Tag becomes part of the garment (not “another thing to manage”). The goal of the Health Tag’s industrial design is to reach adherence levels matching that of undergarments. The Health Tag uses fabric adhesive to attach permanently at specified areas: front of inner underwear waistband (Figure 1a) and inner wing of the bra for females (Figure 1b).

By adhering to garments instead of skin, health tags circumvent issues associated with skin-adhering patches and electrodes such as skin irritation and maintenance repositioning. Health tag patients are not required to adopt new behavioral habits such as remembering to put the device on and having their health device visible to others. Health tags may contain any number of on-board sensors.

Each Health Tag contains a battery, tri-axis accelerometer, Bluetooth radio, vibro-tactile feedback mechanism, photoplethysmography sensor, and force sensor to sense relative changes in abdominal and/or thoracic excursion. The present study quantitatively assesses the accuracy of the respiratory effort monitoring when worn in this manner.

IV. METHOD

A laboratory study was performed in which participants wore both a gold standard respiratory sensing apparatus (herein “ground truth”) and two Health Tags (males: two on waist; females: one on waist, one on bra) while performing a battery of tasks. We aimed to compare minute-by-minute respiratory rate and the temporal location of the four primary components of each respiratory cycle (inspiration-begin, inspiration-end, expiration-begin, and expiration-end). This latter group of metrics was chosen because many other respiratory metrics can be deduced from this set [12].

25 participants (12 female; mean age: 29.8 (SD=7.1); mean body mass index (BMI): 24.0 (SD=3.6)) were told the nature of the study and to dress accordingly during recruitment. The study was conducted in accordance with the principles outlined in the Declaration of Helsinki and approved by the Institutional Review Board of Solutions IRB.

A. Ground Truth

Unlike relying on capnography alone (which employs only nasal cannula), full sensing of respiration requires capture across both mouth and nose. This is relevant when considering the myriad ways in which the mouth is used for respiration (e.g., speech, coughing). After testing multiple portable metabolic analyzers, none guaranteed a constant sample rate with no data loss required for true time synchronization. Thus, an FDA-approved CPAP mask (AirFit, Resmed Corp., San Diego, CA) was connected directly with an FDA-approved digital air flow meter (SFM3000, Sensirion Corp, Staefa, Switzerland) used in hospital ventilatory applications. The test sensor was sampled at 25 Hz, ground truth at 250 Hz, then both time synchronized at millisecond granularity.

B. Data Extraction

The four temporal markers were identified automatically in the ground truth (smoothed with a low pass filter) signal as defined below.

The markers were defined in terms of air flow rather than force (i.e., positive values reflect inspiration while negative values reflect expiration). Inspiration-begin was the first...
positive deflection of the flow signal after negative deflection. Inspiration-end was the return to the zero-line from a positive deflection. Expiration-begin was the the first negative deflection after the zero-crossing. Expiration-end is the return to the zero-line from a negative deflection. A human expert then performed quality assurance testing on the automated scoring on all runs on the truth signal. The same marker identification algorithm was run on the smoothed test signal.

Ground truth RR was calculated after collection using a 2-stage process: first, the number of whole respiratory cycles was counted within a 1-minute window using the zero-crossing algorithm. Second, the time over which those cycles occurred was re-measured by taking the expiration-end point (the trace crossing back over the zero crossing from a trough) of the last cycle in the window and subtracting the time of the inspiration-begin of the first cycle in the window. Finally, the number of cycles in the trimmed segment was counted to assess respiratory rate (breaths per minute, bpm).

**TABLE I. STUDY PROTOCOL**

<table>
<thead>
<tr>
<th>Timestamps</th>
<th>Task</th>
<th>Posture</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00 – 20:00</td>
<td>Rest</td>
<td>Seated</td>
<td>Watch a nature video</td>
</tr>
<tr>
<td>21:00 – 30:00</td>
<td>Desk work</td>
<td>Seated</td>
<td>Trivia using laptop</td>
</tr>
<tr>
<td>31:00 – 40:00</td>
<td>Desk work</td>
<td>Stand</td>
<td>Continue previous task</td>
</tr>
<tr>
<td>41:00 – 50:00</td>
<td>Speech</td>
<td>Seated</td>
<td>Reading aloud from a book</td>
</tr>
<tr>
<td>51:00 – 60:00</td>
<td>Stress</td>
<td>Seated</td>
<td>See [15]</td>
</tr>
<tr>
<td>61:00 – 69:00</td>
<td>Walking</td>
<td>Treadmill</td>
<td>Comfortable pace</td>
</tr>
<tr>
<td>70:00 – 78:00</td>
<td>Running</td>
<td>Treadmill</td>
<td>Strenuous pace</td>
</tr>
<tr>
<td>79:00 – 87:00</td>
<td>Cycling</td>
<td>Ergometer</td>
<td>Medium pace</td>
</tr>
</tbody>
</table>

C. Study Protocol

After a 20 minute intake and sensor fitting period, participants began the protocol in Table 1. In between each task there was a 1-minute standing rest period. The desk work tasks had implied time pressure in that they were presented as a list to ‘get through’, the goal being to simulate office work. The facilitator monitored the Bluetooth connection in real-time to ensure a steady connection with the test sensors.

D. Statistical Analysis

Per [23], error relative to the gold standard was calculated for RR using the formula:

\[
\text{Error} = \frac{\text{device_measurement} - \text{ground_truth}}{\text{ground_truth}}
\]

Aggregate median error, mean absolute error, and standard deviation of the error were computed per sensor position and task type. The task types were physical (walk, run, cycle) and cognitive (all remaining tasks). A Pearson product-moment correlation coefficient was computed to assess the correlation between the minute-by-minute respiratory rates assessed using ground truth and test sensors. An \( r \) value of < 0.70 was considered to be not correlated. Finally, to investigate agreement between the two signals more closely, a Bland–Altman analysis was performed for all runs and the results plotted with 95% confidence intervals (Figure 5).

V. RESULTS AND DISCUSSION

25 subjects doing 8 tasks led to 200 total task runs. For each run analyzed, the entirety of the task duration was used (i.e., not only the final minute of the task). Data from 20 of the 200 total runs (from 7 participants) was discarded because participants asked to remove the mask during those runs due to physical discomfort. The runs where this occurred were noted by the study administrator and, in post hoc analysis, data from those specific runs was removed. Participants almost always resumed use of the mask after a brief rest period and their subsequent data was included in the analysis.

Table 2 contains the median absolute percentage error, mean absolute error (MAE) and standard deviation, root mean squared error (RMSE), and Pearson’s correlation coefficient for each sensor location and task type. The overall error observed across all runs was 6.8%. The highest error was observed during physical activity tasks with test devices worn at the waist; the lowest error was observed using sensors worn at the chest during physical tasks.

**TABLE II. MINUTE-BY-MINUTE RESPIRATORY RATE ACROSS SENSOR POSITION AND TASK TYPE BETWEEN TEST AND GROUND TRUTH**

<table>
<thead>
<tr>
<th>Sensor Position</th>
<th>Task Type</th>
<th>Median Error</th>
<th>MAE±SD</th>
<th>RMSE</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>All</td>
<td>6.8%</td>
<td>1.8±0.14</td>
<td>2.7</td>
<td>0.89</td>
</tr>
<tr>
<td>Waist</td>
<td>All</td>
<td>7.7%</td>
<td>2.1±0.17</td>
<td>3.2</td>
<td>0.94</td>
</tr>
<tr>
<td>Chest</td>
<td>All</td>
<td>5.4%</td>
<td>1.2±0.07</td>
<td>1.2</td>
<td>0.96</td>
</tr>
<tr>
<td>All</td>
<td>Cog</td>
<td>6.9%</td>
<td>1.4±0.13</td>
<td>2.0</td>
<td>0.88</td>
</tr>
<tr>
<td>Waist</td>
<td>Cog</td>
<td>7.6%</td>
<td>1.6±0.16</td>
<td>2.3</td>
<td>0.79</td>
</tr>
<tr>
<td>Chest</td>
<td>Cog</td>
<td>5.7%</td>
<td>1.1±0.07</td>
<td>1.5</td>
<td>0.94</td>
</tr>
<tr>
<td>All</td>
<td>Phys</td>
<td>6.7%</td>
<td>2.5±0.16</td>
<td>3.8</td>
<td>0.65</td>
</tr>
<tr>
<td>Waist</td>
<td>Phys</td>
<td>8.8%</td>
<td>3.2±0.19</td>
<td>4.8</td>
<td>0.47</td>
</tr>
<tr>
<td>Chest</td>
<td>Phys</td>
<td>4.6%</td>
<td>1.5±0.06</td>
<td>2.1</td>
<td>0.89</td>
</tr>
</tbody>
</table>

\( a \) Hypothesis testing indicated that the data from the test signal in these conditions were not significantly different from those derived from ground truth.

\( b \) At task completion, subjects were informed the speech component was not required.
Physical activity monitoring at the waist was the only condition not highly correlated ($r<0.70$). The strongest relationship was found when the test device was worn at the chest. The weakest relationship was observed on waist-worn test sensors during physical tasks. We attributed this to physical exertion causing thoracic breathing (as opposed to abdominal), where it could be sensed more readily by the chest-worn sensor.

Figure 4 depicts the minute-by-minute respiratory rates of ground truth (x-axis) and test sensors (y-axis) across all runs. The Pearson product-moment correlation resulted in a 89.2% correlation between the two signals overall ($r=0.892, p<0.001$).

Figure 5 depicts the results of the Bland-Altman analysis computed on the estimated and ground truth RR values, which yielding a mean bias is -1.04 bpm and limits of agreement at -2.63 and 1.60 bpm. Systematic bias was small; the measures indicate symmetry around the mean bias line.

Figure 6 illustrates the median error across the different sensor locations and task types. Using the figure, we observe that sensor performance across worn positions and task types is generally consistent, with median absolute percentage error consistently around 6%. A second look indicates that sensor performance on the chest is more consistent and less varied than waist-worn sensors. Further, the median error generally falls under 10% across all tasks and sensor positions.
Analysis of the respiratory cycle markers were done by first computing the temporal distance of each estimated marker from its corresponding ground truth marker. Table 3 displays descriptive statistics about these differences. To put these values into context, the average adult respiratory cycle is 3-5 seconds long. To simplify this further, one could consider a 1:1 inspiration-to-expiration ratio, leaving both components at 1.5-2.5 seconds. Using this table, we see the standard deviations are consistent across marker type, indicating that all markers have a similar level of dispersion within them.

Figure 7 illustrates the frequency with which the temporal distance was binned for each of the markers. This figure illustrates how all markers are front-heavy towards greater followed by a tail.

These figures and results indicate a generally high level of precision of the temporal markers when compared with ground truth. The largest deviation (0.40 seconds mean samples difference) is observed for expiration-end. We attribute this to the mechanistic nature of expiration itself: compared to inspiration, the slope of expiration is gradual and its end point is difficult to determine. Inspiration is a fairly rapid event resulting in a sudden deflection and followed by expiration almost immediately. This is clinically relevant, for example, in training respiratory muscles to wean off ventilators.

Figure 8 illustrates the deviation of the Health Tag-assessed temporal markers from the ground truth after outliers (2.5 * interquartile range) are removed. The median difference is consistently under 0.5 seconds, even for expiration-end.

The present study is an observational study of the accuracy of the Health Tag device. The most significant clinical implication is leveraging longitudinal respiratory data. The novel form factor theoretically allows for adherence and utility far exceeding existing wearable respiratory monitors.

A second implication is for the types of data that can be deduced from the Health Tag data. To this end, the respiratory rate results across all sensor locations and types of tasks indicate error is with well-cited precedent [17] though there is no definitive standard. While clinically relevant range of errors are not known for the respiratory temporal markers, the results illustrate that clinically-relevant biomarkers based on such markers could theoretically be reliably reported.

### Study Limitations

The study protocol was short (87 minutes) and does not include supine postures. Conducting a longitudinal in situ study is a matter for future work. Second, the discomfort of the mask caused a number of runs to be discarded from the analysis, and though these did not occur in a particular task, the number was significant. Third, as mentioned above, the CPAP mask was not comfortable to use and thus may have affected natural respiration. The mask was chosen to ensure respiration from both nose and mouth was captured in the truth signal. Finally, the purported intent of health tags is longitudinal data capture but the study in question was not longitudinal in nature.

### Table III. Accuracy of Temporal Markers of Major Respiratory Cycle Components

<table>
<thead>
<tr>
<th>Temporal Marker</th>
<th>Mean Sample Diff.</th>
<th>Std Dev</th>
<th>Median Sample Diff.</th>
<th>Mean Abs. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inspiration Begin</td>
<td>0.66</td>
<td>0.87</td>
<td>0.32</td>
<td>0.24</td>
</tr>
<tr>
<td>Inspiration End</td>
<td>0.65</td>
<td>0.89</td>
<td>0.28</td>
<td>0.20</td>
</tr>
<tr>
<td>Expiration Begin</td>
<td>0.64</td>
<td>0.89</td>
<td>0.26</td>
<td>0.22</td>
</tr>
<tr>
<td>Expiration End</td>
<td>0.72</td>
<td>0.91</td>
<td>0.40</td>
<td>0.28</td>
</tr>
</tbody>
</table>

All but SD are in seconds units.

Fig. 8. Box plots illustrating the estimated difference from ground truth of Health Tag assessments of temporal markers of the major components of the respiratory cycle. IB=inspiration begin, IE=inspiration end, EB=expiration begin, EE=expiration end.
VI. CONCLUSION
Clinical decision-making has long suffered from a lack of continuous, longitudinal, and direct monitoring of respiratory effort. The present study introduces the health tag, a novel form factor attached to the patient’s existing clothing. This paper first introduces a novel form factor that theoretically addresses the factors that cripple device adherence. The paper then validates the respiratory data of a commercial device built in this form. The results indicate that, across both respiratory rate and more granular temporal markers, the device in question produces data within a clinically acceptable range of error across both cognitive and physical activity. The greatest error is observed during physical activity when the sensor is worn at the waist. This is the first known validated technique to monitor respiratory effort continuously at this level of granularity while theoretically reducing patient burden to afford longitudinal capture and analysis.

A. Disclaimers
The authors of this paper are employees of Spire, Inc.

REFERENCES
[7] Elsarnaawy, T; Farrag, Manal; Haueisen, Jens; Abulaal, Magdy; Mahmoud, Khalid; Fouda, H.; Ansari, S. G. A Wearable Wireless Respiration Rate Monitoring System Based on Fiber Optic Sensors. Sensor Letters, Volume 12, Number 9, September 2014, pp. 1331-1336(6)