This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

Objective Scoring of Physiologically Induced Dyspnea by Non-invasive RF Sensors

Zijing Zhang*, Pragya Sharma, Thomas Bradley Conroy, Veerawat Phongtankuel and Edwin C. Kan, Senior Member, IEEE

***Abstract*— *Objective:* Dyspnea, also known as the patient's feeling of difficult or labored breathing, is one of the most common symptoms for respiratory disorders. Dyspnea is usually self-reported by patients using, for example, the Borg scale from 0 – 10, which is however subjective and problematic for those who refuse to cooperate or cannot communicate. The objective of this paper was to develop a learning-based model that can evaluate the correlation between the self-report Borg score and the respiratory metrics for dyspnea induced by exertion and increased airway resistance. *Methods:* A non-invasive wearable radio-frequency sensor by near-field coherent sensing was employed to retrieve continuous respiratory data with user comfort and convenience. Self-report dyspnea scores and respiratory features were collected on 32 healthy participants going through various physical and breathing exercises. A machine learning model based on the decision tree and random forest then produced an objective dyspnea score. *Results:* For unseen data as well as unseen participants, the objective dyspnea score can be in reasonable agreement with the self-report score, and the importance factor of each respiratory metrics can be assessed. *Conclusion:* An objective dyspnea score can potentially complement or substitute the self-report for physiologically induced dyspnea. *Significance:* The method can potentially formulate a baseline for clinical dyspnea assessment and help caregivers track dyspnea continuously, especially for patients who cannot report themselves.**

***Index Terms*—Dyspnea; respiration sensors; medical diagnosis.**

I. Introduction

THE symptom of dyspnea, or so called difficult or labored breathing, defined as a "patient's subjective awareness of uncomfortable or distressing breathing" [1], can be caused by heavy exertion, deficiency of ambient oxygen, increased airway resistance, and respiratory disorders. Dyspnea can be extremely distressing for patients with serious illness [2], such as asthma, heart failure, COVID-19 [3], and chronic obstructive pulmonary diseases (COPD) [4], [5], [6], leading to a poor quality of life [7]. The prevalence of dyspnea is common, and ranges from 33% to 76% in critically ill patients, while 85% of patients with heart failure and up to 95% of patients with COPD report dyspnea. In addition to the possible physiological consequence in cardiopulmonary functions such as low oxygen saturation levels in SaO2 and SpO2 [8], the repercussion of dyspnea can also include the sense of suffocation, distress, fear, panic or anxiety. Although dyspnea has been associated with the intertwined physiological, psychological, and social-demographic contributors [9], [10], no present theory was able to encompass all causes of dyspnea reliably. Our study design focuses on the physiological factor of dyspnea that can be induced from exertion and increased airway resistance, and we hypothesize that this kind of dyspnea will have a high correlation between the self-report and the measurable respiratory features.

In present clinical practices, dyspnea is most often assessed by different kinds of scales self-reported by patients [11]. The most popular scales include the following scores conducted during patient interviews for specific purposes.

1) The perceived disability scale, such as the Medical Research Council (MRC) dyspnea scale, describes the breathlessness sensation out of the exercise capacity in a score of 1 – 5 [12], where Score 1 denotes that the participant is not troubled by dyspnea except on strenuous exercise, and Score 5 is for patients who are too breathless to leave the house or during dressing/undressing.

2) The experiential history scale, such as the baseline and transition dyspnea index (BDI/TDI), measures changes in the three domains of dyspnea severity from functional impairment, magnitude of task, and magnitude of effort [13]. In BDI, the scale is from 0 – 12, which is the sum of 0 – 4 in each of the three domains of dyspnea severity. Score 0 denotes the severe

This work is supported by Department of Defense of United States through Office of the Congressionally Directed Medical Research Programs (CDMRP) Discovery Award PR-182496, and by National Institute of Health (NIH) R21 DA049566-01A1.

Z. Zhang, P. Sharma, T. B. Conroy, and E. C. Kan are with School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA. V. Phongtankuel is with Geriatrics and Palliative Medicine, Weill Cornell Medical College, New York, NY 10065, USA. E-mails: {zz58, ps847, tbc38, eck5}@cornell.edu; vep9012@med.cornell.edu).

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

dyspnea and Score 4 for unimpaired functions. The TDI scale ranges from –3 (major deterioration) to +3 (major improvement) in each domain from previous BDI. The popular use of BDI/TDI also illustrates the importance of continuous evaluation of dyspnea episodes.

3) The psychophysical scale, such as the Borg scale or visual analog scale (VAS), assesses symptom intensity in response to a specific stimulus such as exercise. In the Borg scale, patients report their feeling of discomfort from 0 – 10 at the moment, where 0 corresponds to the sensation of normal breathing or absence of dyspnea, and 10 corresponds to the maximum possible sensation of dyspnea [14], [15]. This is similar to the other popular Borg scale for pain.

Regardless of the scale of choice, self-reported dyspnea is subjective and variable for each person on each day, and can be challenging to assess for those who refuse to cooperate or cannot communicate due to medical issues such as stroke, dementia, and delirium. Frequent quires to patients for continuous dyspnea evaluation are not only tedious, but can also cause stress and discontent, introducing a bias to the self-reported score. To supplement the subjective patient self-report, the association between dyspnea and respiratory metrics, such as the breath rate (BR), lung volume (LV), and inhalation/ exhalation patterns, has been carefully examined in previous works [16]. Studies of the association have been however cross sectional, lacking the ability to account for continuous metrics and their changes, often due to the discomfort or inconvenience of sensor deployment on patients [17]. Despite these limitations in past studies, it is of critical importance to ameliorate the objective methods of dyspnea evaluation, which can help clinicians make better decision on treatment and triage in patients with chronic lung and heart diseases, can enhance screening effectiveness in pandemic situations, and can enable caregivers in palliative and hospice medicine to provide timely service to those patients who are unable to communicate due to cognitive impairment, loss of language ability, or delirium.

However, current methods of respiration monitoring [18] such as respiratory inductance plethysmography (RIP)[19], strain gauge (SG) [20], spirometer [21], pneumotachometer [22], and capnography [23] are seldom practical for use outside of clinical settings due to the need of operator assistance and patient cooperation. Many of these devices are uncomfortable or require long-time connection to immobile machines, which are thus not feasible for broad deployment, especially on chronical patients with advanced lung diseases. RIP and tension-based chest belts can measure respiratory waveforms from the thoracic and abdominal movement. However, the required tight belt tension to respond to the full waveform inevitably causes discomfort. Body movement to relax the belts to the least constraining position also brings additional artifacts. The SG sensor requires tight skin contact which is also uncomfortable for long-term monitoring. Spirometry can measure the volume and speed of inhaled and exhaled air, but involves attentive user participation and strenuous breathing protocol. The pneumotachometer is the gold-standard device for measuring airflow by detecting the pressure drop against a small resistive field, but is cumbersome to deploy and uncomfortable for users. Capnography measures the amount of CO_2 in exhaled air to derive the respiratory effort and distress, but the device is bulky and patients need to be intubated. Radar-based sensing [24], [25], [26] as another alternative respiration monitoring method can assess body surface motion as a result of breathing, but requires a reader in the direct line of sight (LoS) to the front chest of the user. Ambient and body motion can cause severe interference, although BR during tidal breathing can be reasonably derived through careful filtering.

In this paper, we opt to use the wearable radio-frequency (RF) sensors [27], [28] based on near-field coherent sensing (NCS) to monitor respiratory features over layers of clothing or furniture fabrics to enhance user comfort and convenience. In the near-field region, the dielectric boundary movement by internal organs and muscles will be modulated on the antenna characteristics, and then be received in a multiplexed channel [28] in a non-invasive manner. Due to its touchless operation, many subjects under test will not notice the on-going sensing activities at all [29].

The sensation of dyspnea can be voluntarily and objectively induced from either physical exertion or increase in airway resistance, which is the assumption behind the MRC scale as the perception of the exertion capabilities. We began our study by collecting data from 32 healthy participants wearing the respiratory sensors including the NCS sensor and the commercial chest belts for sensor validation. Participants were instructed to report dyspnea scores D_{self} frequently in the Borg scale under various exertion and breathing exercises following a given protocol. Scales based on patient perception or experience such as MRC and TDI were less feasible for real-time exertion-based studies. Four respiratory features of BR, LV, and inhalation and exhalation intervals were extracted from the continuous waveforms. We defined 15 respiratory metrics out of the 4 respiratory features, and constructed the machine-learning (ML) models based on the decision tree [30] and random forest [31] to investigate the correlation between the respiratory metrics and D_{self} during the entire protocol. The resulting ML model can produce an objective dyspnea score D_{obj} in cross validation with D_{self}, and can also identify the individual importance factors of respiratory metrics in determining D_{obj}.

The main contribution of this paper is to provide a new way to generate the objective score for physiologically induced dyspnea, using a comfortable and continuous respiratory sensor and an established ML model which can simultaneously consider multiple factors with different importance weighting. As far as we know, this is the first study to examine the association between dyspnea sensation and continuous respiratory metrics that account for changes in respiratory behavior over a period of time under exertion and increased airway resistance. The objective dyspnea score D_{obj} can potentially complement or substitute the self-report dyspnea score D_{self}. When more comprehensive clinical data in established patient population are available for training in the future, this model can potentially assist clinicians and caregivers in more reliable diagnosis and treatment of dyspnea.

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

II. Methods

A. Experimental setup

NCS is based on the near-field coupling of ultra-high frequency (UHF) electromagnetic (EM) waves with the dielectric boundary movement of internal organs and body parts. Fig. 1(a) shows the NCS deployment together with the reference commercial sensors by BIOPAC (BIOPAC Systems, Inc., Goleta, CA), including electrocardiogram (ECG) and tension-based torso belts. Fig. 1(b) shows the photo of two software-defined radios (SDR) by National Instrument Ettus B210 to implement the NCS transceivers [28], [29]. We utilized both chair-integrated notched transmission lines [32] and wearable antennas [27] as the NCS sensing elements, as shown in Figs. 1(c)(d). For the wearable sensor, we used an antenna pair as the transmitter (Tx) and receiver (Rx). UHF radiation can penetrate dielectrics in the near-field of the sensor, so the dielectric composition will modulate the EM distribution, and be exhibited in the cross-coupling scattering parameter S_{21} of the antenna pair. Alternatively, the notched sensor was constructed by a miniature coaxial RF cable with the metal shield of the middle part removed. One end will be connected to Tx and the other to Rx. The dielectric boundary movement by lungs and associated muscles will be coupled to the leaked EM energy from the notched part of the RF cable, and hence can be detected either at Tx as signal reflection S_{11} or at Rx as signal transmission S_{21}. Notice that the SDR setup can use a digitally modulated superheterodyne signal to improve the channel isolation from ambient RF signals [27], [32], and high-quality measurements do not rely on anechoic RF chambers.

The two wearable NCS sensors were placed on the chest and the abdomen in the front torso, and the two notched sensors were integrated to the back of a chair behind the thorax and abdomen. The two sets of NCS sensors can be operated individually or at the same time. A participant wearing all sensors, including NCS, the torso belts and ECG in the sitting posture, is shown in Fig. 1(d) when breathing through an N95 facemask. BIOPAC and two sets of NCS sensors were included to verify that the conclusion on association between D_{obj} and D_{self} would be sensor independent. The exercise bike used in the exertion protocol is shown in Fig. 1(e). ECG is used to provide supplementary heartbeat information. The synchronization between SDR and BIOPAC recordings is achieved by buffering both data at approximately the same time within a few milliseconds.

Both SDR and BIOPAC were connected to the host computer through USB (Universal Serial Bus). In the SDR Tx signal chain, the digital baseband went through the digital-to-analog converter (DAC) and was then mixed with the carrier frequency f_{RF}. The RF power is less than −10 dBm or 0.1 mW, well under the safety limits set by occupational safety and health administration (OSHA) in the UHF band. The RF signal emitted from Tx was then coupled into the internal organ motion within the near field. The modulated signal was received by Rx, and then demodulated and sampled by the analog-to-digital converter (ADC) to retrieve the motion characteristics in the baseband. The quadrature baseband signals $I_{TX}(t)$ and $Q_{TX}(t)$ were presently implemented by sinusoidal monotones with the NCS signal as amplitude modulation [27]:

$$I_{Rx}(t) = NCS(t)\cos(2\pi f_{BB}t + \theta_0) \quad (1)$$

$$Q_{Rx}(t) = NCS(t)\sin(2\pi f_{BB}t + \theta_0) \quad (2)$$

$$NCS_{amp}(t) = \sqrt{I_{Rx}(t)^2 + Q_{Rx}(t)^2} \quad (3)$$

where θ_0 was the phase offset accumulated from the Tx-Rx signal chains. f_{RF} was selected at 900 MHz for the wearable sensors, and 1GHz for the notched sensors. Each B210 was used for two Tx-Rx channels with the two basebands at f_{BB1} =355 kHz and f_{BB2} = 440 kHz, both sampled at 1M samples per second (Sps). The demodulated respiration waveform was further down-sampled to 500 Sps, which was the same as the BIOPAC data.

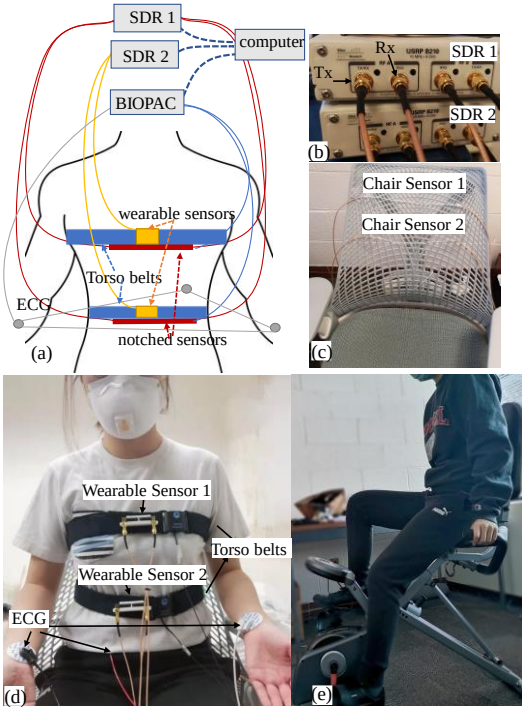


Fig. 1. The experimental setup: (a) Schematics of NCS and BIOPAC sensors and data flow; (b) The photo of software-defined radio (SDR) transceivers that were connected to NCS sensors;(c) Chair-integrated NCS sensor setup; (d) A participant wearing NCS and BIOPAC torso belt sensors in the sitting position during breathing through a facemask. Two NCS sensors and two belts were deployed at the thorax and abdomen position. (e) A participant undergoing physical exertion. Written informed consent was obtained from the participants to publish their photos.

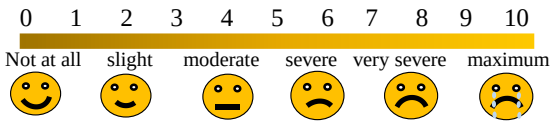
B. Human Study Protocol

The experimental setup described in Sec. II.A was applied to evaluating the respiratory waveforms of 32 healthy participants whose demographic distribution is shown in Supplementary Table I. The human study has been approved by Cornell Institutional Review Board (IRB) Protocol ID #1812008488. Written informed consent to take part in the study was obtained from all participants. Participants were instructed to follow a sequence of routines as documented in Table I. The breathing exercise includes: 0 – 30s fast breathing, 30 – 60s slow

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

breathing, 60 – 90s normal breathing, 90 – 120s fast breathing, 120 – 150s slow breathing, and 150 – 180s normal breathing to build a library of various breathing patterns. The participant followed a voice instruction at the beginning of each section, and reported a dyspnea score D_{self} in the Borg VAS scale after each routine as shown in Fig. 2 [14]. The participant sat on a chair for all routines except during Routines 4 and 6 of exertion. Dyspnea was induced by aerobic rope jumping and exercise biking, as well as by wearing an N95 mask to increase the airway resistance. Fig. 3 presents examples of NCS respiratory waveforms during the study protocol. In comparison, BIOPAC torso-belt waveforms are presented in Supplementary Fig. 1. When the routines in dyspnea (red curves) were compared with routines without dyspnea (green curves), in Routine 3, BR decreases due to partial airway obstruction and breath-to-breath variation also decreases; in Routines 5 and 7, BR increases and breath-to-breath variation decreases due to post-exertion. We will further analyze these changes by extracting quantitative metrics in the respiratory waveforms in the later sections.



0: Relaxed as in sleeping or watching TV.
1-3: Can maintain for hours.
4-6: Can hold short conversation. But noticeable challenging.
7-9: Difficult to maintain exercise. Can barely speak a sentence.
10: Impossible to keep going. Unable to talk.

Fig. 2. Description of the self-reported Borg visual analog scale (VAS) for dyspnea evaluation [14].

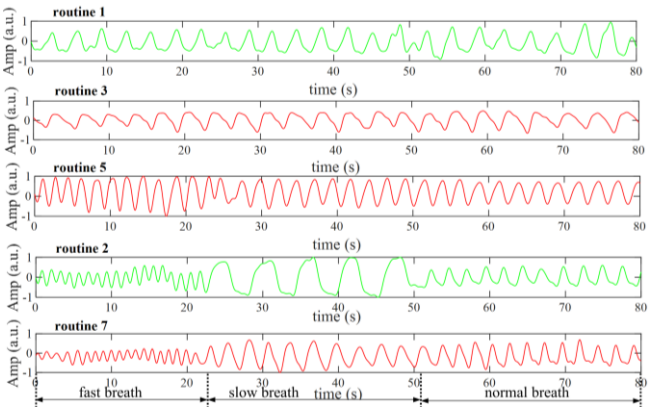


Fig. 3. An example of NCS respiratory waveforms during the study protocol. Routine 1: Normal breathing; Routine 3: Normal breathing with a facemask; Routine 5: Normal breathing after physical exertion; Routine 2: Breathing exercise; Routine 7: Breathing exercise after physical exertion. The illustration period is truncated from 10 to 90 s of each routine. Green curves indicate absence of dyspnea, and red curves indicate some degrees of dyspnea.

III. DATA PROCESSING

A. *Physiological manifestation of dyspnea*

The main purpose of respiration is to supply oxygen to body cells by circulation, with the auxiliary functions of making sound, sniffling, and clearing of airway by coughing and sneezing. Respiration can be initiated involuntarily and voluntarily, and the voluntary part can be trained. When the blood oxygen saturation (SaO2) is low or CO_2 high, the breathing action will be triggered for more lung ventilation. However, when the body cannot respond fast enough due to various reasons such as airway obstruction, insufficient ambient oxygen supply, weakened respiratory muscles, or voluntary control for speaking, singing or holding, the feeling of dyspnea will arise quickly. To increase lung ventilation, often BR and LV will increase by panting or deep breathing. Alternatively, the inhalation and exhalation intervals will be adjusted depending on the muscle condition, airway obstruction, and ambient factors. As the respiratory reaction to dyspnea can be trained to reduce the uncomfortable feeling, similar to experiential avoidance coping of pain, another common physiological reaction to dyspnea is the reduction of variability in successive breaths [16] together with speaking restraint, when the body tries hard to use the best known breathing cycle to reduce the discomfort of dyspnea.

Therefore, we propose to use the respiratory features of BR, LV, and inhalation and exhalation intervals to correlate to dyspnea manifestation. The mean and variation of these features within a chosen epoch as well as the variation between successive breaths will be extracted from the breathing waveforms for further data processing. Notice that here we will not complicate our protocol with speaking and coughing, as they can be separately identified from their high-frequency characteristics [33]. For future extension to realistic continuous monitoring, both speaking and coughing will need to be accounted for.

B. *Feature extraction*

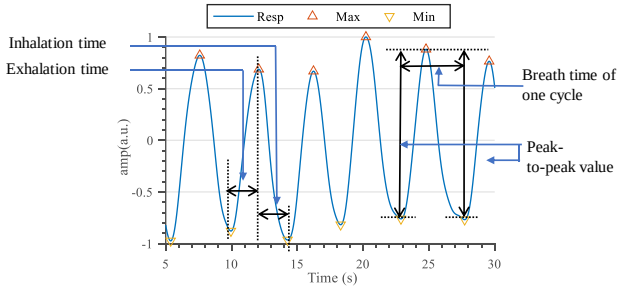


Fig. 4. An example of feature extraction from the NCS respiratory waveforms. The blue line labeled as 'resp' is the raw NCS waveform after bandpass filtering of 0.05 – 1 Hz, and the red and yellow triangles are detected peaks of maximum and minimum by the moving-average crossing method. The breath rate, peak-to-peak value, inhalation interval and exhalation interval can be estimated out of the detected peaks.

The retrieved respiratory waveforms from 4 NCS and 2 BIOPAC torso-belt sensors were first bandpass-filtered from 0.05Hz to 1Hz to remove the DC drift and high-frequency noises. Various sensor combinations will be further studied in the dyspnea recognition below. The filtering processing was implemented in MATLAB by the digital infinite impulse response (IIR). We then utilized the moving average-crossing algorithm [34] to detect peaks of the breathing waveform. A moving-average curve was first calculated at each time point in a given window length, which was around one respiration cycle

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

and would be constantly updated. The points when the moving-average curve crossed the original signal were marked as up-crossing points for positive slopes in the original signal or down-crossing points for negative. Local maximum was labelled as the maximal point between two up-down crossing points, and local minimum as the minimal point between two down-up crossing points. As shown for an example in Fig. 4, the blue line was the filtered respiratory waveform from the wearable NCS sensor, and the red triangles and yellow triangles marked the maximum and minimum peaks detected by the algorithm. Then, we can extract the 4 respiratory features in each breath cycle to represent the instantaneous respiratory characteristics: 1) BR was calculated from the inverse of the interval between two neighboring minima; 2) The peak-to-peak (PP) value representing LV [27] was estimated by the signal difference in successive peaks; 3) The inhalation interval (IN) was evaluated by the time difference between one minimum and the following maximum; 4) The exhalation interval (EX) between one maximum and the following minimum. The respiratory waveform from the torso belts was processed in the same way. The peak-detection algorithm is of critical importance for accurate feature extraction and subsequent processing.

C. *Respiratory metrics*

TABLE I. RESPIRATORY METRICS FROM RESPIRATORY FEATURES.

	Breath rate (BPM)	Peak-to-peak (a.u.)	Inhalation interval (s)	Exhalation interval (s)
Coefficient of variation	CoV_{BR}	CoV_{PP}	CoV_{IN}	CoV_{EX}
Mean	μ_{BR}		μ_{IN}	μ_{EX}
Autocorrelation	$R1_{BR}$	$R1_{PP}$	$R1_{IN}$	$R1_{EX}$
Successive differences	$R2_{BR}$	$R2_{PP}$	$R2_{IN}$	$R2_{EX}$

After calculating the above 4 respiratory features, we defined 15 metrics that serve as the input to the ML model as shown in Table I. The first 4 metrics were the coefficient of variation (*CoV*) of the above 4 respiratory features, which was defined as

$$CoV = \left(\frac{\sigma}{\mu}\right)^2 \quad (4)$$

where σ denotes the standard deviation and μ denotes the mean. At each sampling point, *CoV* was calculated over an epoch of T = 15s. We first found all breath cycles in the previous 15s and calculated BR of every cycle, whose μ_{BR} and σ_{BR} were estimated to obtain CoV_{BR}. CoV_{BR} were then averaged over all epochs in the given routine. *CoV* for PP, IN and EX were derived in the same way. These respiratory metrics can represent the breath variability within the epoch. Mean values were estimated as μ_{BR}, μ_{IN} and μ_{EX} for the BR, inhalation and exhalation intervals during each routine. As PP was normalized and contained the bias from personal deployment, μ_{PP} was excluded from the respiratory metrics.

To further capture variability between adjacent breaths, we used autocorrelation in a time lag of one breath cycle to measure the successive similarity of a given respiratory feature. The autocorrelation function *R1* is defined as,

$$c_i = \frac{1}{N}\sum_{n=1}^{N-1} y_n \cdot y_{n+i} \quad (5)$$

$$R1 = \frac{c_1}{c_0} \quad (6)$$

where n is the breath index within a total of N breaths in the epoch of T = 15s during metrics evaluation, y_n is the discrete breath-by-breath measurement of the selected respiratory feature, and y_{n+i} is the same feature lagged by i breaths. We also define a similarity measure of *R2* as the mean absolute difference between adjacent breaths:

$$R2 = \frac{1}{N}\sum_{n=1}^{N-1} \frac{|y_{n+1} - y_n|}{y_n} \quad (7)$$

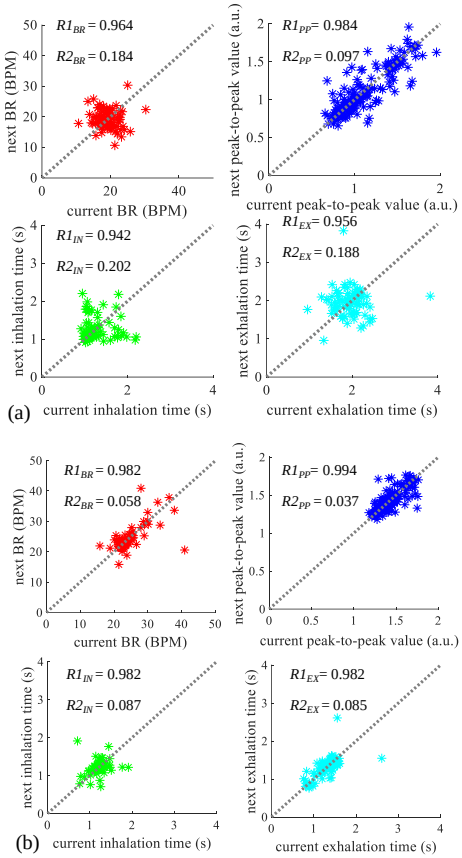


Fig. 5. Analysis of the breath rate (BR), lung volume (LV) by peak-to-peak magnitude, inhalation interval and exhalation interval between adjacent breaths by NCS during 5-min recordings. (a) Normal breathing under a self-report dyspnea score of 0; (b) Normal breathing after exercises with a self-report dyspnea score of 7. Decrease of variability in the respiratory features can be observed by increasing R1 and decreasing R2.

Fig. 5 shows a representative sample analysis using the NCS recording for two routines. Fig. 5(a) is from the routine with D_{self} = 0, i.e., no sense of dyspnea, while Fig. 5(b) from a routine with D_{self} = 7 after heavy exercises. For the higher D_{self}, μ_{BR} is higher but CoV_{BR} remains similar, μ_{PP} is higher with smaller CoV_{PP}, and μ_{IN}, CoV_{IN}, μ_{EX} and CoV_{EX} are all reduced due to the faster breathing with more regularity. Increase in autocorrelation R1 and decrease in successive differences R2 are consistently observed for all 4 respiratory features in our data, implicating reduction in variability between adjacent breaths.

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

D. Sensor fusion considerations

To investigate the error from the sensor hardware, we compare the analyses using the wearable and chair-integrated NCS as well as the BIOPAC chest belts. To expand the dyspnea model, we also added the cardiogram features from NCS bandpass-filtered between 0.5 and 3 Hz or BIOPAC ECG to the combined model learning. Three heartbeat metrics were included: 1) The mean heart rate (HR) μ_{HR}, 2) the standard deviation of NN intervals σ_{NN} [35], where NN denotes the normal RR distance in the QRS cardiogram complex, and 3) the root mean square of the successive differences between adjacent NN RMS_{NN}. When both NCS and BIOPAC data were used together, the number of cases would be doubled. The heartbeat signals added 3 additional metrics to the original 15 metrics of respiration. Notice that ideally the NCS and BIOPAC sensors would measure the same cardiopulmonary features, but each had its own noises and errors during measurements that can cause inaccurate prediction of D_{obj}. Combination of sensors measuring the same physiological features may or may not improve the overall accuracy, as inconsistent derivation of intended metrics can aggravate the ambiguity, unless the noise can be assumed to be totally uncorrelated in the fusion of a large number of sensors. The combination of two sets of NCS and BIOPAC sensors cannot guarantee such an assumption. To understand whether the wearable or chair-integrated NCS sensor can be used alone in the applications, we compare each individual data set and the combination as well.

IV. RESULTS

A. Personal calibration

We extracted the respiratory and heartbeat metrics from the breathing and cardiac waveforms of 32 healthy participants using a protocol with various levels of exertion and airway obstruction in Table II. The breathing exercises in Routines 2 and 7 contain specified inhalation and exhalation instructions for the participant to follow, which cover multiple periods of fast, slow and hold breathing. Each routine in the protocol contains a self-reported dyspnea score D_{self} and all measured respiratory metrics. Because the tidal breathing pattern varied from person to person, we opted to first calibrate out the personal difference by subtracting the respiratory metrics in the normal breathing routine (D_{self} = 0 as the baseline) from the exertion routines (after exercise or wearing a facemask with D_{self} = 1 – 9) for the same participant.

TABLE II. ROUTINES IN THE HUMAN STUDY PROTOCOL.

Routine	Duration (minutes)	Routine content	Respiratory monitoring
1	5	Normal breathing	ON
2	3	Breathing exercises	ON
3	5	Normal breathing with facemask	ON
4	3 OR 10	Rope jumping or exercise biking	OFF
5	5	Normal breathing	ON
6	3 OR 10	Rope jumping or exercise biking	OFF
7	5	Breathing exercise	ON

This personal calibration against the normal breathing routine reduced the individual biases in D_{obj}, but the case of D_{obj}= 0 would be excluded from the ML model output. It is important that the eventual model can also give reliable D_{obj}= 0 for negative dyspnea cases. To remedy the case of D_{self} = 0, we imputed the input with the new cases where the respiratory metrics in the first half of normal breathing was subtracted from the second half. The difference between the two halves of the first normal breathing routine was used as the training under D_{self} = 0. Extreme dyspnea scenario (D_{self} = 10) was not included in this paper due to safety concerns in the human study. Extrapolated respiratory metrics for complete obstructive and central apnea have been attempted by data imputation, but with only limited success. Hence, the extreme cases of D_{self} = 10 will be left for future clinical studies when experimental observation can be available.

B. The decision-tree and random-forest models

To produce D_{obj} from the measured respiratory metrics, we chose the decision-tree regressor [30] as the ML model for the following reasons. (1) Decision tree is a white-box model, so the physical explanation for the result can be observable through the tree structure, which can help us understand the physiological correlation between the respiratory metrics and D_{self}. (2) Decision tree helps dominant feature selection in multitudinous respiratory metrics. Irrelevant respiratory metrics will be assigned a less importance weight to evolve with the dominant features, and the importance factors can be part of the model output for physiological reasoning. (3) The regressor model is preferred over the classifier because the output can be a continuous quantity of the predicted D_{obj}. Although D_{self} in the Borg scale is discrete for subjective convenience, a continuous D_{obj} can reduce the ambiguity between fine discrete levels. To illustrate the advantage of decision-tree regressor against the popular discrete classification methods such as the principal component analysis (PCA), we showed the scattered plots of the two dominant features of D_{self} found by the decision-tree model in Supplementary Fig. 2. It can be observed that clusters of D_{self} of similar values cannot be identified in any reasonable hyperplane separation. As PCA can only capture linear correlation, the inability of classification by PCA indicates that the correlation between D_{self} and respiratory metrics is more convoluted and ambiguous. In comparison, the decision tree can incorporate nonlinear relationship into the model with reasonable tolerance of ambiguous contributors.

TABLE III. DATA COMPOSITION IN THE ML MODEL.

Training set (95 cases):		
Number of cases	D_{self}	Respiratory metrics
32	1 – 9	Calibrated normal breathing after exercise in Routine 5
31	1 – 9	Calibrated normal breathing with facemasks in Routine 3
32	0	Imputed normal breathing by two halves of Routine 1
Testing set (30 cases):		
23	1 – 9	Calibrated breathing exercise after exercise in Routine 7
7	0	Imputed normal breathing by two halves of Routine 1

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

TABLE I
PREDICTION ACCURACY FROM DATA SETS IN TABLE III.

Data set	NCS + torso-belt + ECG	NCS	torso-belt
Model	Decision tree	Decision tree	Decision tree
Mean η by k-fold cross-validation	0.876	0.825	0.818
Feature importance	μ_{BR}=0.402 $R2_{IN}$=0.149 CoV_{Ex}=0.076 RMS_{NN} =0.068	μ_{BR}=0.402 $R2_{BR}$=0.151 $R2_{EX}$=0.075 $R1_{IN}$ =0.071	μ_{BR}=0.330 $R2_{IN}$=0.290 $R2_{BR}$=0.078 CoV_{Ex}=0.074
Mean η by leave-one-participant-out cross-validation	0.841	0.864	0.854
Feature importance	μ_{BR}=0.463 CoV_{Ex}=0.126 $R2_{IN}$=0.111 RMS_{NN} =0.083	μ_{BR}=0.443 $R2_{BR}$=0.198 CoV_{PP}=0.054 $R2_{EX}$=0.045	μ_{BR}=0.330 $R2_{IN}$=0.262 CoV_{Ex}=0.074 $R2_{BR}$=0.056
η for testing data	0.872	0.871	0.834
Data set	NCS + torso-belt + ECG	NCS	torso-belt
Model	Random forest	Random forest	Random forest
Mean η by k-fold cross-validation	0.884	0.866	0.848
Feature importance	μ_{BR}=0.280 $R2_{IN}$=0.150 $R2_{BR}$=0.103 CoV_{Ex}=0.054	μ_{BR}=0.232 $R2_{BR}$=0.164 $R2_{EX}$=0.103 $R1_{IN}$ =0.082	μ_{BR}=0.223 $R2_{IN}$=0.184 $R2_{BR}$=0.130 CoV_{BR}=0.068
Mean η by leave-one-participant-out cross-validation	0.874	0.881	0.866
Feature importance	μ_{BR}=0.332 $R2_{IN}$=0.161 $R2_{BR}$=0.068 CoV_{Ex}=0.056	μ_{BR}=0.232 $R2_{BR}$=0.164 $R1_{IN}$=0.106 $R2_{EX}$=0.080	μ_{BR}=0.229 $R2_{IN}$=0.183 $R2_{BR}$=0.128 CoV_{Ex}=0.087
η for testing data	0.903	0.907	0.873

After training in the 95 cases in Table III, the decision-tree model can predict D_{obj} on the unseen 30 testing cases based solely on the respiratory metrics, which can then be compared with D_{self} to assess the model accuracy. In Table III, the 95 training cases consist of measurements from Routine 1, 3 and 5 with observation after exertion and during facemask wearing, as well as cases of D_{self} = 0 for calibration and imputation. One participant opted out of the facemask wearing routine during the study.

In order to estimate the skill of our model on new data, we first performed the procedure of k-fold and leave-one-participant-out cross-validations on the 95 training cases [36]. K-fold cross-validation can investigate the robustness to unseen data, and leave-one-participant-out cross-validation can test the robustness to unseen participants. For k-fold cross-validation, we divided the whole training set of 95 cases into separate training (76 cases) and testing (19 cases). We chose k = 5 and the model was trained using 4 folds as the training data and the resulting model is validated on the remaining fold as the testing data. For leave-one-participant-out cross-validation, the model was trained on the data sets from 31 participants excluding one participant, who was then used as testing by generating D_{obj} to compare with D_{self}. The validation process was repeated for each participant as the testing case.

Because we used the regressing estimator, the predicted D_{obj} can be a continuous number from 0 to 9. The upper limit of D_{obj} = 9 is due to the lack of training cases with D_{self} = 10. We define a prediction accuracy η of the ML model as:

$$\eta = 1 - \frac{|D_{obj} - D_{self}|}{9} \quad (8)$$

where the error is the normalized absolute distance of D_{obj} to D_{self}. The maximum value of η for perfect prediction is 1. In the k-fold cross-validation procedure, mean η for the validation set ranges from 0.818 to 0.881. In the leave-one-participant-out cross validation, mean η fell within similar ranges. When testing on the 30 unseen data set as shown in Table IV, η ranges from 0.834 to 0.907. The 30 unseen testing set contains Routine 7 recording from 23 participants, as 9 participants out of 32 cannot follow the breathing exercise after exertion, as well as 7 participants who had repeated Routine 1 in different study dates. The accuracy and the metrics of importance for all three testing scenarios are summarized in Table IV. The metric with a higher importance factor has a higher correlation with D_{self}. The sum of the importance factors from all features was normalized to 1 in each method. To understand the magnitude of mean η better, we performed total random guesses of D_{obj} for the unseen 30 testing cases in Table III, where η would range from 0.566 to 0.677. A fixed guess of D_{obj} in all dyspnea prediction will render η ranging from 0.396 (D_{obj} = 9) to 0.766 (D_{obj} = 4). The fixed-guess η for all D_{obj} is summarized in Supplementary Fig. 3 on the 30 unseen testing cases. When more cases are available with homogeneous distribution across all possible values of D_{self}, η will approach 0.5 for random and fixed guesses.

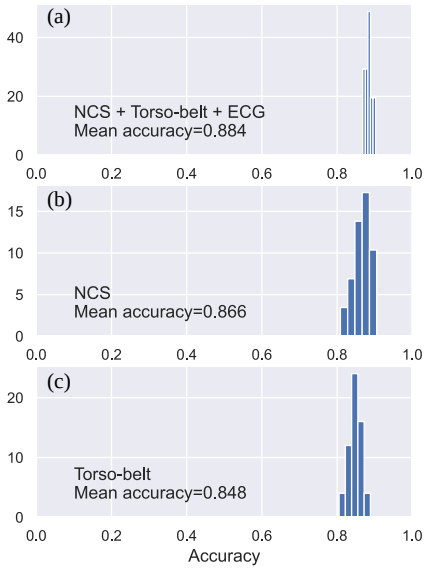


Fig. 6. Accuracy distribution by k-fold cross-validation using the random forest model. (a) Training data from wearable NCS + torso-belt +ECG; (b) Wearable NCS only; (c) BIOPAC torso-belt only.

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

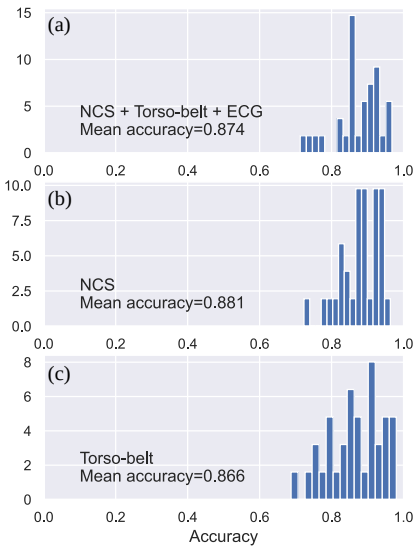


Fig. 7. Accuracy distribution by leave-one-participant-out cross-validation using the random forest model. (a) Training data from wearable NCS + torso-belt + ECG; (b) Wearable NCS only; (c) BIOPAC torso-belt only.

In Table IV, we added the random-forest method composed by an ensemble of decision trees [31] as the ML model. We first used the *k*-fold cross-validation method to choose multiple decision trees that have high prediction accuracy in the 95 training sets, and then their average would be used to predict D_{obj} for the 30 unseen testing data. We also listed the top four most important metrics in generating the final random-forest model under consideration. As observed from Table IV, our model can predict the dyspnea score in the unseen testing data with reasonable accuracy. Random forest can only improve η marginally, but tends to have smaller variation, as will be seen in the later Bland-Altman (B&A) plots. This was possibly because the set of optimal decision trees was reasonably similar, as can be observed from the dominant feature importance with the same input. The accuracy difference between NCS and NCS + BIOPAC was also trivial, which suggested that the NCS sensor has captured the respiratory metrics with sufficient accuracy, and the torso belts in BIOPAC do not offer new information.

Fig. 6 presents the accuracy distribution during *k*-fold cross-validation using the random forest model. Fig. 6(a) used the training data from NCS + torso-belt + ECG sensors, Fig. 6(b) from only NCS sensors, and Fig. 6(c) from only torso-belt sensors. Fig. 7 presents the similar accuracy distribution by leave-one-participant-out cross-validation. By separating one independent participant's data as the testing set and estimating the accuracy on each participant, our model has been evaluated to have a good performance to predict an unseen participant. As a special case of the *k*-fold cross-validation when the number of folds equals the number of instances in the data set, the leave-one-participant-out cross-validation can offer lower bias but manifest higher variance over some participants when their body types or the cooperation levels differ from the rest of the group. It can also tend toward model overfitting because the learning and testing data sets are strongly imbalanced during the validation process, and because the learning iterations contain major overlaps in the training data. In general, NCS performs slightly better than the torso belt in cross validation, but the difference is probably within error ranges. Notice that the two sensors have similar signal-to-noise ratios (SNR) [27]. SNR estimation by the spectral analysis in our test data is summarized in Supplementary Fig. 4.

Additional scattered plots for all columns in Table IV are shown in the Fig. 8(a)(b) where the error in different D_{self} values can be more clearly observed. We also investigated the effectiveness of the wearable and chair-integrated NCS in Table V, as these two setups can be applied to different clinical applications, for example, the wearable sensor for patients in the pulmonary ward and the chair-integrated sensor in the observation room. The accuracy difference between wearable and chair-integrated NCS was insignificant. We can also conclude that the touchless NCS sensors alone can generate the ML model with high validity in either the wearable or chair setup. Scattered and B&A plots for Table V are shown in Figs. 8(c)(d) and 9, where NCS in both setups are shown to produce D_{obj} with reasonable limits of agreement (LoA) and no significant systematic bias *m*. The random-forest model has tighter LoA than the decision-tree model, probably due to the variation reduction during ensemble averaging.

TABLE V. COMPARISON OF NCS WEARABLE AND CHAIR SETUPS.

Data set	Wearable NCS	Chair NCS	Wearable NCS	Chair NCS
Model	Decision tree	Decision tree	Random forest	Random forest
Mean η by *k*-fold cross-validation	0.822	0.825	0.865	0.866
Mean η by leave-one-participant-out cross-validation	0.861	0.865	0.883	0.883
η for unseen testing data	0.855	0.879	0.906	0.903

TABLE VI. TESTING RESULTS BY SEPARATING EXERTION AND FACEMASK.

Data set	NCS + torso-belt + ECG	NCS	Torso-belt
Model	Decision tree	Decision tree	Decision tree
Mean η by *k*-fold cross-validation	0.937	0.868	0.900
Mean η by leave-one-participant-out cross-validation	0.940	0.917	0.941
η for unseen data	0.798	0.773	0.764
Data set	NCS + torso-belt + ECG	NCS	torso-belt
Model	Random forest	Random forest	Random forest
Mean η by *k*-fold cross-validation	0.944	0.878	0.902
Mean η by leave-one-participant cross-validation	0.939	0.905	0.929
η for unseen data	0.802	0.828	0.790

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

To further test the transferrable capability of the ML model, we also shifted the 31 cases of normal breathing with facemask in Routine 2 from the training set to the testing set. Facemask only increases the lung elasticity, and causes the dyspnea sensation in a different manner from exertion. The result was presented in Table VI. The accuracy for the cross validation becomes slightly higher as the data were collected from more consistent routines. The accuracy for the unseen data from Routine 2 was lower than those in Table IV, partially because less training data was provided and the testing data came from totally unseen routines. However, the accuracy remains reasonably high, which showed the model extendibility to predict dyspnea scores in unseen routines of different dyspnea contributors, possibly even in patients with various respiratory disorders. The corresponding scattered plots for the columns of the unseen data in Table VI are shown in Figs. 8 (e)(f), where both torso-belt and NCS had non-negligible numbers of cases where D_{self} was high but D_{obj} was close to 0. As the dyspnea after exertion is used for training and the dyspnea during facemask wearing for testing, these cases indicated that the respiratory features in Routine 2 of facemask wearing was more similar to normal breathing and less to dyspneic episodes after exertion, even though the participant reported a high D_{self}. However, D_{obj} from 'NCS + torso-belt + ECG' had much fewer such cases. Our conjecture is that the cardiac information can help during model transference.

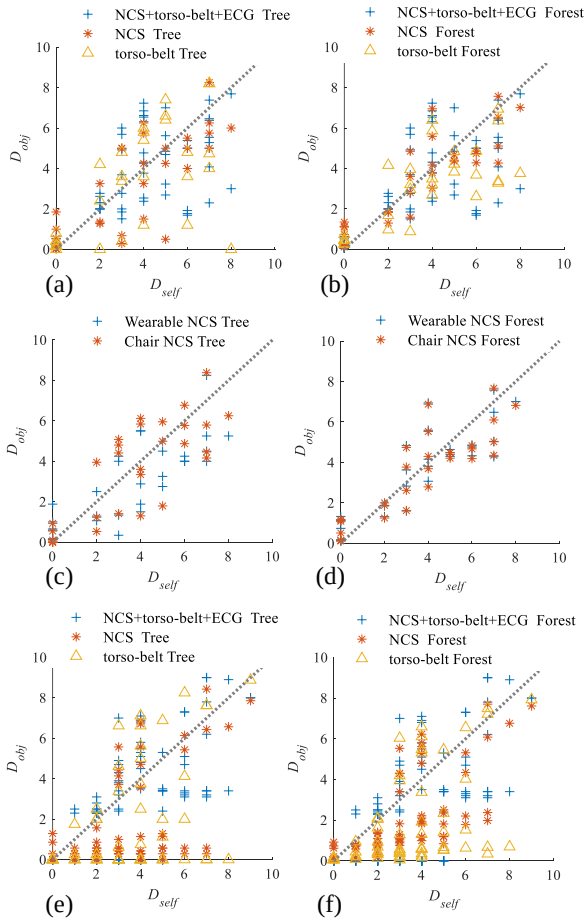


Fig. 8. The scatter plots between D_{self} and D_{obj} for all columns in (a) (b) Table IV, (c) (d) able V, and (e) (f) Table VI.

Further observation from Table IV indicates that the additional information from ECG boosts the prediction accuracy in general, although not by much, especially in the case of the random-forest model. RMS_{NN}, a form of the heart-rate variation (HRV), is the fourth most important metric for the sensor fusion with ECG, although the importance factor is much smaller than the successive breath variation. During dyspnea sensation, people are often under some degrees of psychological stress, which can then indirectly influence HRV [37], [38] to become a contributor for dyspnea recognition. For the two cases in Table VI where the increased airway resistance is in the testing data but not in the training data, ECG can sometime causes overall accuracy degradation, as in the case of random forest between 'NCS + torso-bet + ECG' and NCS alone. This is also explainable as mask wearing can cause certain degrees of stress for some people who had to make efforts to breath in sufficient air. However, during the COVID-19 pandemic era when the human study had been done, many people were used to wearing masks without any physical or psychological stress. Thus, the heartbeat information becomes an inconsistent and ambiguous contributor in the facemask wearing cases that can degrade the dyspnea prediction.

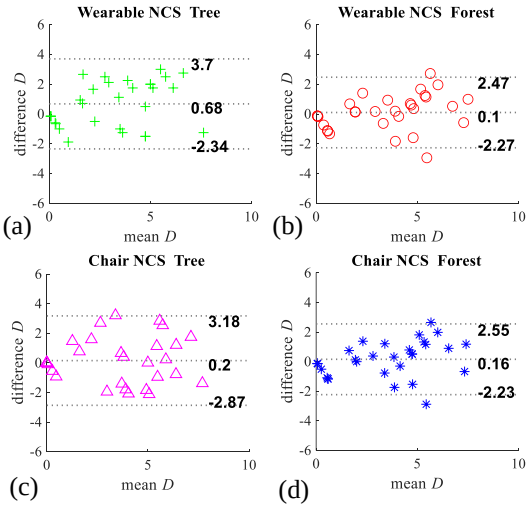


Fig. 9. The Bland-Altman (B&A) plots between D_{self} and D_{obj} for all columns in Table V. Plots show the mean difference m at the center dotted line and the corresponding limits of agreement (LoA) at the upper and lower dotted lines given by $m \pm 1.96\sigma$.

V. Limitations

In this paper, barriers to accomplish a mature D_{obj} model still remain. The present limitations are listed in the following. 1) D_{self} has individual biases and tolerances due to the multiple factors in the cause and outcome of dyspnea even for the same person in different days of testing, so the training data might not be repeatable and inevitably contained subjective variations, which will be trickling down to the training model as ambiguity. Nevertheless, the Borg scale has been designed with small intervals, and small variations in D_{self} can be acceptable from the large data set. 2) To establish the simplified baseline of objective dyspnea scoring, our present study only includes healthy participants with physiologically induced dyspnea by exertion and facemasks in a prescribed study protocol. The

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

respiratory features by psychological contributors or in patients with chronic cardiopulmonary diseases have not been examined, and can be different from dyspnea induced by physical exertion. Clinical studies should be extended to demonstrate that respiratory waveforms from patient dyspnea have similar features as those from this study. 3) The signal processing and ML models employed in this study are previously established with limited novelty and some respiratory features that correlate to dyspnea detection had been reported elsewhere. However, instead of proposing new ML models that may carry further uncertainties or biases, we have carefully chosen the white-box decision-tree and random-forest models that can not only give accurate dyspnea representation, but also provide insights to the physiological reasoning.

VI. Discussion

In the future, it will be critical to extend to large-scale patient studies on real patients with various cardiopulmonary disorders to establish the true effectiveness of the proposed ML model. Despite of its limitations, the present study still offers a useful starting point for the future studies of specific respiratory disorders, especially towards the patient population who refuse to cooperate or cannot communicate due to deterioration or loss of mental functions, where the D_{obj} model has to be achieved through transference as individual D_{self} cannot be available.

VII. Conclusion

In conclusion, dyspnea is an important common symptom for respiratory disorder screening, diagnosis and prognosis. Objective dyspnea scores to complement self-report are especially important for patients with serious illness and at the end of life when communication or cooperation is compromised due to dementia, delirium, anesthesia, and other restraining procedures. In this paper, we show that the non-invasive NCS sensors can continuously capture useful respiratory features during various levels of dyspnea physiologically induced by exertion and increased airway resistance. We have performed human study on 32 healthy participants and constructed a learning-based model to identify the correlation between continuous respiratory metrics and self-reported dyspnea score, and hence can predict an objective dyspnea score induced by physiological reason with reasonable accuracy. In future clinical study, there can be additional intertwined contributors to dyspnea in patients under different disorders and conditions, but our present study can provide a baseline of physiological analyses and a useful reference to the eventual prediction model.

References

[1] M. L. Campbell, "Respiratory distress: A model of responses and behaviors to an asphyxial threat for patients who are unable to self-report," *Heart & Lung*, vol. 37, no. 1, pp. 54-60, Jan. 2008.

[2] A. H. Kamal *et al.*, "Dyspnea review for the palliative care professional: Assessment, burdens, and etiologies," *J. Palliat. Med.*, vol. 14, no. 10, pp. 1167-1172, Oct. 2011.

[3] D. Wang *et al.*, "Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China," *Jama*, vol. 323, no. 11, pp. 1061-1069, Mar. 2020.

[4] J. P. Solano, B. Gomes, and I. J. Higginson, "A comparison of symptom prevalence in far advanced cancer, AIDS, heart disease, chronic obstructive pulmonary disease and renal disease," *J. Pain Symptom Manag.*, vol. 31, no. 1, pp. 58-69, June 2006.

[5] D. C. Currow *et al.*, "A community population survey of prevalence and severity of dyspnea in adults," *J. Pain Symptom Manag.*, vol. 38, no. 4, pp. 533-545, 2009.

[6] D. C. Currow *et al.*, "Do the trajectories of dyspnea differ in prevalence and intensity by diagnosis at the end of life? A consecutive cohort study," *J. Pain Symptom Manag.*, vol. 39, no. 4, pp. 680-690, Apr. 2010.

[7] N. Karnani, G. Reisfield, and G. R. Wilson, "Evaluation of chronic dyspnea," *Am. Fam. Physician*, vol. 71, no. 8, pp. 1529-1537, Apr. 2005.

[8] M. Tsao, E. Barnes, and E. Chow, "The relationship between dyspnea and blood oxygen saturation," *J. Pain Symptom Manag.*, vol. 12, no. 2, pp. 109-114, 2019.

[9] K. Tanaka *et al.*, "Factors correlated with dyspnea in advanced lung cancer patients: Organic causes and what else?," . *J. Pain Symptom Manag.*, vol. 23, no. 6, pp. 490-500, June 2002.

[10] T. Janssens *et al.*, "Dyspnea perception in COPD: Association between anxiety, dyspnea-related fear, and dyspnea in a pulmonary rehabilitation program," *Chest*, vol. 140, no. 3, pp. 618-625, Sep. 2011.

[11] E. Crisafulli and E. M. Clini, "Measures of dyspnea in pulmonary rehabilitation," *Multidiscip. Respir. Med.*, vol. 5, no. 3, p. 202, Jun. 2010.

[12] J. Bestall *et al.*, "Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease," *Thorax*, vol. 54, no. 7, pp. 581-586, July 1999.

[13] D. A. Mahler *et al.*, "The measurement of dyspnea: Contents, interobserver agreement, and physiologic correlates of two new clinical indexes," *Chest*, vol. 85, no. 6, pp. 751-758, June 1984.

[14] R. C. Boshuizen, A. D. Vincent, and M. M. van den Heuvel, "Comparison of modified Borg scale and visual analog scale dyspnea scores in predicting re-intervention after drainage of malignant pleural effusion," *Supportive Care in Cancer*, vol. 21, no. 11, pp. 3109-3116, July 2013.

[15] G. A. Borg, "Psychophysical bases of perceived exertion," *Med. Sci. Sports Exerc.*, vol. 14, no. 5, pp. 377-381, 1982.

[16] T. Brack, A. Jubran, and M. J. Tobin, "Dyspnea and decreased variability of breathing in patients with restrictive lung disease," *Am. J. Respir. Crit. Care Med.*, vol. 165, no. 9, pp. 1260-1264, May. 2002.

[17] N. Wolkove *et al.*, "The relationship between pulmonary function and dyspnea in obstructive lung disease," *Chest*, vol. 96, no. 6, pp. 1247-1251, Dec. 1989.

[18] F. Q. AL-Khalidi *et al.*, "Respiration rate monitoring methods: A review," *Pediatr. Pulmonol.*, vol. 46, no. 6, pp. 523-529, Jan. 2011.

[19] G. Brullmann *et al.*, "Respiratory monitoring by inductive plethysmography in unrestrained subjects using position sensor-adjusted calibration," *Respiration*, vol. 79, no. 2, pp. 112-20, Apr. 2010, doi: 10.1159/000212117.

[20] M. Chu *et al.*, "Respiration rate and volume measurements using wearable strain sensors," *NPJ Digit. Med.*, vol. 2, p. 8, Feb. 2019, doi: 10.1038/s41746-019-0083-3.

[21] G. T. Ferguson *et al.*, "Office spirometry for lung health assessment in adults: A consensus statement from the National Lung Health Education Program," *Chest*, vol. 117, no. 4, pp. 1146-1161, Apr. 2000.

[22] M. P. Yeh *et al.*, "Computerized determination of pneumotachometer characteristics using a calibrated syringe," *J. Appl. Physiol.*, vol. 53, no. 1, pp. 280-285, July 1982.

[23] K. Bhavani-Shankar *et al.*, "Terminology and the current limitations of time capnography: A brief review," *J. Clin. Monit.*, vol. 11, no. 3, pp. 175-182, May 1995.

[24] C. Li *et al.*, "A review on recent advances in Doppler radar sensors for noncontact healthcare monitoring," *IEEE Trans. Microw. Theory. Tech.*, vol. 61, no. 5, pp. 2046-2060, May 2013.

[25] W. Massagram, V. M. Lubecke, and O. Boric-Lubecke, "Microwave non-invasive sensing of respiratory tidal volume," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009, pp. 4832-4835.

[26] C. Gu et al., "Accurate respiration measurement using DC-coupled continuous-wave radar sensor for motion-adaptive cancer radiotherapy," *IEEE. Trans. Biomed. Eng.*, vol. 59, no. 11, pp. 3117-3123, June 2012.

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3096462, IEEE Transactions on Biomedical Engineering

[27] P. Sharma *et al.*, "Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate," *NPJ Digit. Med.*, vol. 3, p. 98, July 2020, doi: 10.1038/s41746-020-0307-6.
[28] X. Hui and E. C. Kan, "Monitoring vital signs over multiplexed radio by near-field coherent sensing," *Nat. Electron.*, vol. 1, no. 1, pp. 74-78, Nov. 2017, doi: 10.1038/s41928-017-0001-0.
[29] X. Hui and E. C. Kan, "No-touch measurements of vital signs in small conscious animals," *Sci. Adv.*, vol. 5, no. 2, art. eaau0169, Feb. 2019.
[30] L. Breiman *et al.*, *Classification and Regression Trees*. Belmont, CA: Wadsworth (International Group). 1984, pp. 151-166.
[31] T. K. Ho, "The random subspace method for constructing decision forests *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832-844, Aug. 1998.
[32] Z. Zhang *et al.*, "Furniture-integrated respiration sensors by notched transmission lines," *IEEE Sens. J.*, vol. 21, no. 4, pp. 5303-5311, Feb. 2021.
[33] P. Sharma, X. Hui, and E. C. Kan, "A wearable RF sensor for monitoring respiratory patterns," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, July 23 – 27, 2019, pp. 1217-1223.
[34] W. Lu *et al.*, "A semi-automatic method for peak and valley detection in free-breathing respiratory waveforms," *Med. Phys.*, vol. 33, no. 10, pp. 3634-6, Oct. 2006, doi: 10.1118/1.2348764.
[35] R. R. Singh, S. Conjeti, and R. Banerjee, "A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 740-754, Nov. 2013.
[36] J.-H. Kim, "Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap," *Computational Statistics and Data Analysis*, vol. 53, no. 11, pp. 3735 – 3745, Sep. 2009.
[37] W. von Rosenberg *et al.*, "Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV," *Front. Physiol.*, vol. 8, p. 360, June 2017.
[38] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Med. Biol. Eng.*, vol. 42, no. 3, pp. 419-427, May 2004.

0018-9294 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Cornell University Library. Downloaded on September 14,2021 at 13:27:40 UTC from IEEE Xplore. Restrictions apply.