Predicting Pulmonary Function from Phone Sensors

Qian Cheng, MS,[1,2] Joshua Juen, MS,[2,3] Shashi Bellam, MD,[4] Nicholas Fulara, MS,[5] Deanna Close, RN,[5] Jonathan C. Silverstein, MD,[6] and Bruce Schatz, PhD[2,7]

[1]*Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, Illinois.*
[2]*Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois.*
[3]*Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois.*
Departments of [4]*Medicine and* [5]*Respiratory Therapy, NorthShore University HealthSystem, Evanston, Illinois.*
[6]*Center for Biomedical Research Informatics, NorthShore University HealthSystem, Evanston, Illinois.*
[7]*Department of Medical Information Science, University of Illinois at Urbana-Champaign, Urbana, Illinois.*

Abstract

Introduction: *Smartphones are ubiquitous, but it is unknown what physiological functions can be monitored at clinical quality. Pulmonary function is a standard measure of health status for cardiopulmonary patients. We have shown phone sensors can accurately measure walking patterns. Here we show that improved classification models can accurately predict pulmonary function, with sole inputs being motion sensors from carried phones.*

Subjects and Methods: *Twenty-five cardiopulmonary patients performed 6-minute walk tests in pulmonary rehabilitation at a regional hospital. They carried smartphones running custom software recording phone motion. Each patient's pulmonary function was measured by spirometry. A universal model, based on support vector machine, then computed the category of function with input from signal processing features and patient demographic features.*

Results: *All but a few of every 10-second interval for every patient was correctly predicted. The trained model perfectly computed the GOLD (Global Initiative for Chronic Obstructive Lung Disease) level 1/2/3, which is a standard classification of pulmonary function. Each level was determined to have a characteristic motion, which could be recognized from the sensor features. In addition, longitudinal changes were detected for 10 patients with multiple walk tests, except for cases with clinical instability.*

Conclusions: *These results are encouraging toward clinical validation of passive monitors running continuously in the background, for patients in homes during daily activities. Initial testing indicates the same high accuracy as with active monitors, for patients in hospitals during walk tests. We expect patients can simply carry their phones during everyday living, while models support automatic prediction of pulmonary function for health monitoring.*

Keywords: *health monitoring, telemedicine, mobile phones, chronic disease assessment, pulmonary function, predictive modeling, machine learning*

Introduction

There is a revolution in health monitoring, due to mobile devices. Individual measurement can generate population cohorts of similar patients with similar status, so treatments can be effectively and efficiently targeted.[1] Mobile phones are ubiquitous in the United States, with the Pew Internet Project showing 90% ownership in 2014, including 64% with smartphones. Even seniors more than 65 years of age have 74% penetration of mobile phones.[2] As hundreds of millions of patients are already carrying phones, opportunity emerges for passive monitoring without adherence difficulties. We seek clinically valid physiological measures through carried smartphones.

Of the many possible measures, the most important for diagnostic purposes is functional status. Pulmonary function is measured with a medical spirometer, which a patient breathes into. The volume exhaled is calibrated to provide standard measures of pulmonary function. We show classification models can accurately compute pulmonary function. In clinical testing with rehabilitation patients, we further show that adequate inputs are the motion sensors already contained

© Qian Cheng et al. 2017; Published by Mary Ann Liebert, Inc. This Open Access article is distributed under the terms of the Creative Commons Attribution Noncommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Table 1. Demographic Information of Patient Groups by Global Initiative for Chronic Obstructive Lung Disease (GOLD) Levels

	GOLD1	GOLD2	GOLD3	Overall
Number of patients (Female)	2 (1)	13 (5)	10 (4)	25 (10)
Age (year)	69 (65–73)	80 (67–95)	72 (55–85)	76 (55–95)
Height (m)	1.68 (1.55–1.80)	1.66 (1.24–1.83)	1.68 (1.55–1.83)	1.67 (1.24–1.83)
Weight (kg)	109.5 (93.0–126.1)	80.3 (54.4–112.0)	81.3 (45.4–118.4)	83.1 (45.4–126.1)

Age, height, and weight are in average (minimum – maximum) format. These demographics with gender are used to adjust the model.

GOLD, Global Initiative for Chronic Obstructive Lung Disease.

in ordinary smartphones. Simply carrying phones in daily living can measure health status.

For chronic heart and lung conditions, walk tests are widely used to assess the severity of the disease, including measures with accelerometer sensors.[3,4] The 6-minute walk test (6MWT) is a standard assessment[5] for chronic obstructive pulmonary disease (COPD) and congestive heart failure (CHF), affecting tens of millions of patients. A 6MWT measures the distance walked in 6 min back and forth over a fixed length walkway. Normal gait requires many systems, including strength, sensation, and coordination, function in an integrated manner, so abnormal gait is diagnostic of many conditions.[6] Gait is the total walking pattern, complete body motion, including swaying as well as stepping.

Previously, we used 6MWT to show that motion features could be measured with phones carried by patients, as sole input to trained models that accurately compute gait speed[7] and oxygen saturation,[8] among physiological measures for health status of chronic conditions. This study extends to computing pulmonary function, based on characteristic motions of health status. We have implemented Android phone software to record walking using phone sensors with accuracy similar to medical devices.[9] We designed three phases of experiments to assess using smartphones to predict pulmonary function, from technical evaluation to passive monitoring. The results indicate that this technique is feasible for field trials with large populations.

Subjects and Methods

We recruited 35 pulmonary patients at NorthShore University Health System, under IRB approval starting November 2014. All chronic patients, going through pulmonary rehabilitation in respiratory therapy in Evanston Hospital, were offered the option of participating in our study. Such patients were provided smartphones for recording their motion during a standard 6MWT, a high-end Samsung Galaxy S5 and a low-end LG Optimus Zone2 carried in fanny pack. The 6MWTs were performed on a 30-m straight walkway in the hospital corridor, with cones at each end. The patient walks back and forth under nurse supervision with distance recorded. All 6MWTs follow the standard American Thoracic Society (ATS) guidelines.[5]

Pulmonary function tests (PFTs) were performed with a spirometer under clinical conditions.[10] In obstructive diseases, such as COPD, the ratio of forced expiratory volume in one second (FEV1) as compared with an age, gender, race, and height-adjusted expected value is a sufficient indicator to measure severity level of the disease, called predicted FEV1%. The Global Initiative for Chronic Obstructive Lung Disease (GOLD) defines a standard for pulmonary severity levels based on the predicted FEV1% values.[11] There are four GOLD stages: GOLD1 (mild), GOLD2 (moderate), GOLD3 (severe), and GOLD4 (more severe).

In this study, 35 patients participated in total. Ten patients were eliminated because of lack of either 6MWT or PFT, including 5 patients who did not perform 6MWT, 3 patients who did not have PFTs, and 2 patients who were GOLD4 stopping multiple times. As shown demographically in *Table 1*, 25 patients had both 6MWT and PFT, with 2 GOLD1 (predicted FEV1% $\geq$ 80), 13 GOLD2 patients (predicted FEV1%: 50–79), and 10 GOLD3 patients (predicted FEV1%: 30–49).

FEATURE SELECTION

After preprocessing, we have obtained good walking data for each patient. We compute input features for training the model by the feature selection approach.[7] A primary parameter to measure gait is cadence, number of strides within a unit of time.[12] We calculated cadence by counting steps in each 10-second interval.

Related studies in motion tracking by wearable devices extract various features from raw acceleration data.[13–15] We selected eight sufficient spatiotemporal gait parameters in both time and frequency domain. In time domain, we selected mean and standard deviation of acceleration to describe general distribution of acceleration sample. In addition, mean crossing rate (MCR), root mean square (RMS), autocorrelation coefficient (AC), and coefficient of variance (CV) were calculated from time series acceleration data. MCR represents ratio of above and below acceleration. RMS is statistical measure on the variation of signal magnitude. AC measures periodic similarity in time domain. CV is normalized measure for dispersion of discrete samples. In frequency domain, we

computed peak frequency (PF) and Shannon entropy. PF represents frequency of greatest magnitude in spectrum. Shannon entropy is expected value of signal information.

Besides the cadence and spatiotemporal gait parameters, demographic information must also be considered in model training, just as spirometers are used to adjust the raw values from patient respiration. Unlike selecting demographic cohort and training models by cohort in our previous research,[7] we keep the four basic demographic parameters: age and sex, height, and weight as factors in the input feature vector. With such vector, we train universal models that can be applied on the general population. Overall, for each 10-sec sample, the input feature vector $\vec{x}$ contains 13 independent features—covering stepping, moving, and demographics.

MODEL ANALYSIS

Support vector machines (SVMs) enable classification by remapping multidimensional vectors into higher dimensional space and determining hyperplanes that separate different classes.[16,17] We applied RBFSVM, radial basis function kernel, to train the SVM model.[18] Patient status is labeled from three GOLD stages. One-against-all strategy is applied to train this three-class classification model,[19] then subject status is computed by 85% threshold for majority voting from all 10-sec samples. We designed three phases to analyze performance of the model. The phases included Phase I, a standard 10-fold cross-validation on training set, whereby all walking sessions are involved in model training; Phase II, evaluating model prediction on the testing set, whereby all walking sessions are not involved in model training; and Phase III, applying the trained model with passive walk on new patients, whereby none of the patients contributed to model training.

Results

PHASE I: MODEL VALIDATION WITH LABELED TRAINING SET

The data set contains 1,234 10-sec walking samples (38 6MWTs) from 25 different patients, including 799 labeled samples from the first 6MWT session of each patient. We train classification models only using all labeled samples. Ten-fold cross-validation is applied to self-validate the model. The overall classification accuracy is 99.00%. In detail, all GOLD1 samples are correctly classified, accuracy for GOLD2 is 99.28% and accuracy for GOLD3 is 98.51%. After obtaining classification results for 10-sec samples, majority voting with 85% threshold is applied for categorizing each patient into target GOLD stage.

There were 25 patients who completed one 6MWT session. The model does perfect computation for patients' pulmonary

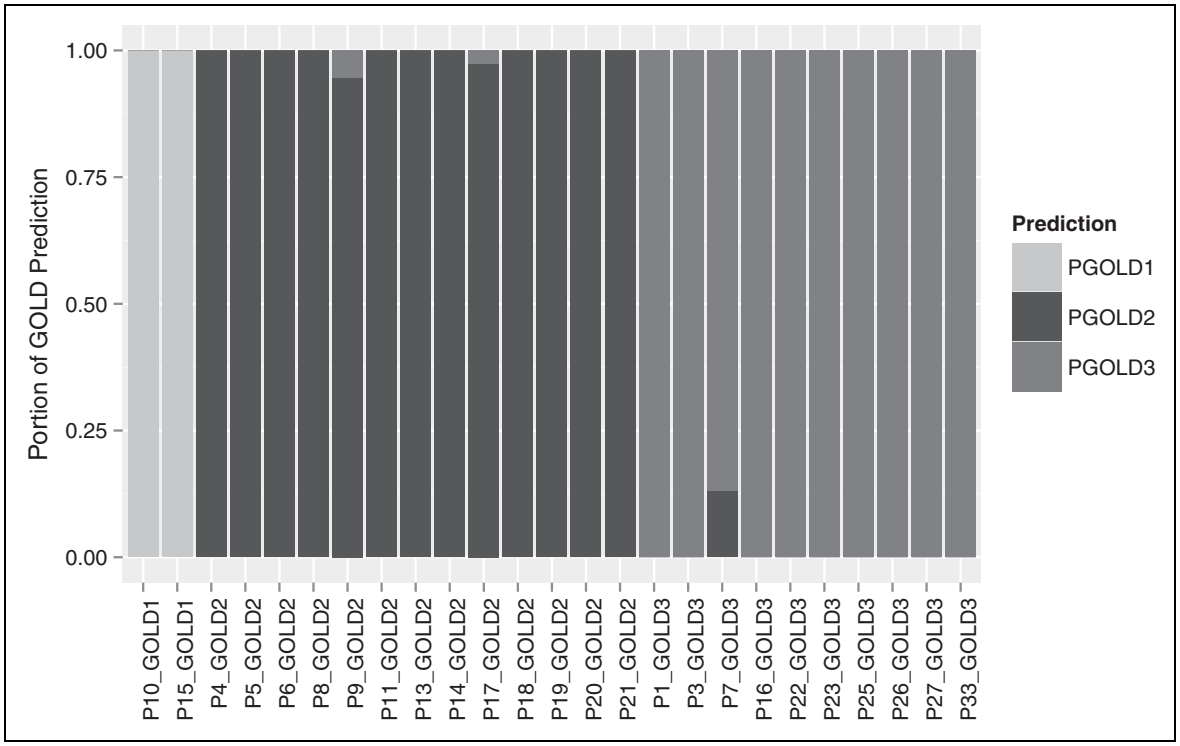

Fig. 1. Evaluation of GOLD classification for labeled 6MWT sessions. GOLD, Global Initiative for Chronic Obstructive Lung Disease; 6MWT, 6-minute walk test.

function, whereby all patients are categorized into the correct GOLD stage. In this case, the PFTs are performed the same day as the 6MWT, so all patients' predicted FEV1% represents their current severity of respiratory limitation when we collect phone motion to match lung function. *Figure 1* shows this computation of health status.

For 22 patients, all their 10-sec walking samples are identically predicted, and correctly reflect the patient's GOLD stage. This indicates that spatiotemporal motion reflects more stable information than walking speed. Patients may speed up or slow down during the 6MWT, but spatiotemporal motion will always yield the correct health status. Thus, full 6MWTs are not necessary for detecting health status of cardiopulmonary patients. Instead, if we can find any good walking sessions identical to the walking during the 6MWT, spatiotemporal motion data from such session can be used to compute GOLD status.

For three patients, a few 10-sec samples are classified incorrectly. For Patient 17, there is only one 10-sec sample with false classification, from 32 eligible samples. For Patient 9, there are two 10-sec samples with false classification, from 37 eligible samples. For Patient 7, 5 samples out of 38 eligible samples get false classification. So voting with 85% threshold does predict correct status. Patient 7 is a GOLD3 patient reflected by the FEV1% predicted value of 31% from his PFT. However, this patient walked for 335 m for this 6MWT, which is even better than general GOLD2 patients. Our model predicted five GOLD2 walk samples, reflecting the patient is healthier than GOLD status during the walk test, matching the evidence of longer walk distance. This patient may be in transition between GOLD2 and GOLD3.

PHASE II: MODEL EVALUATION WITH UNLABELED TESTING SET

We train a classification model with all 799 labeled samples and apply this model on the testing set, forming by the unlabeled samples. That is, the training set is the samples from the first walking session, whereas the testing set is the samples from the subsequent walking session(s), possibly including status change for the patients. There are 435 unlabeled samples from 13 unique 6MWT sessions from 10 patients, whereby 3 patients have three 6MWT sessions (5, 16, and 21). Two of 13 walking sessions are classified differently from the patients' first labeled 6MWT session. For all other patients, classifications of unlabeled 6MWT sessions are 100% identical to their first 6MWT session for all samples, as shown in *Figure 2*.

Six-minute walk distance (6MWD) is considered as a gold standard for diagnosis of COPD severity, but it is more accurate for severe COPD than moderate or mild COPD.[20] In our experiment, 10 patients performed at least two 6MWT sessions, usually several weeks apart. There are 13 sessions, because three

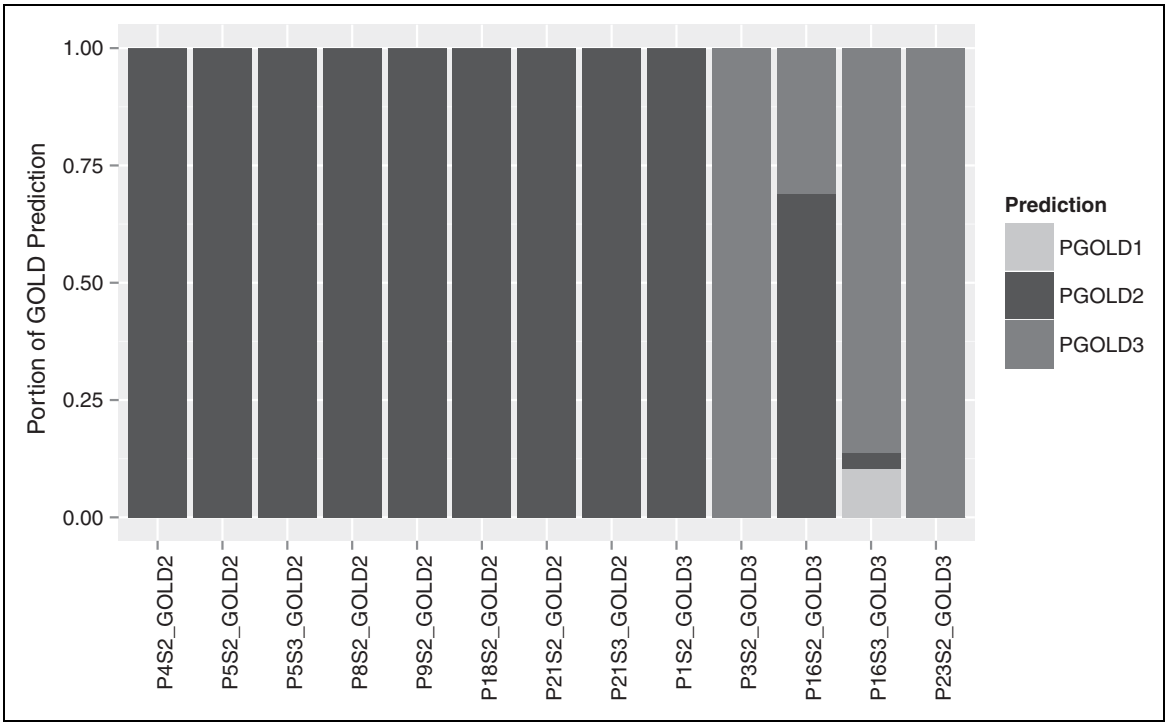

Fig. 2. Evaluation of GOLD classification for unlabeled 6MWT sessions.

patients also did a Session 3. The initial 6MWDs average 310.8 ± 95.16 m and the following 6MWDs average 367.1 ± 106.5 m. In general, the patients walked further in later sessions, as expected for patients undergoing rehabilitation.

Our analysis indicates that 6MWD is not sufficient for measuring COPD severity levels. In our data set, GOLD2 and GOLD3 patients are not well distinguished by 6MWD. The 6MWD of GOLD3 patients can be as high as 350 m but 6MWD of GOLD2 patients can be less than 250 m. Note that variation between two different 6MWDs becomes larger when the distance is shorter, for example, second 6MWD of GOLD3 Patient 23 is 150 m longer than the initial 6MWD, even though performed just 20 days after beginning the rehabilitation.

Our classification model again had high accuracy, because such variation on total distance does not affect the stability of patients' motion pattern. Eight out of 10 patients are detected strictly as having same GOLD status for their second or third 6MWT sessions. All 10-sec walking samples of their following 6MWT sessions are identical to their original GOLD status, even for Patient 23, with the largest variation of 6MWDs. This "unfair" 8 of 10 model accuracy for later sessions compared with initial GOLD status is shown in *Figure 2*.

Two patients not correctly classified at subsequent sessions have underlying clinical explanations. For Patient 1, the second 6MWT session is 100% predicted as GOLD2, whereas the first 6MWT session shows the patient is GOLD3, as FEV1 indicates with spirometry. He had clinical diagnosis of both COPD and CHF when enrolled in pulmonary rehabilitation. His medication regimen for CHF was adjusted at start of rehabilitation and his motion improved, likely from better control of his CHF, so this "unfair" comparison of initial GOLD status to improved GOLD status with the model may be measuring actual improvement rather than model inaccuracy.

For Patient 16, the second 6MWT session is 68.97% predicted as GOLD2 and 31.03% predicted as GOLD3, whereas the first 6MWT session is labeled GOLD3. Patient 16 also has a third 6MWT, which is 10.34% predicted as GOLD1, 3.45% predicted as GOLD2, and 86.21% predicted as GOLD3. So the third session of Patient 16 is decided as GOLD3, same as the first session.

Patient 16 had diagnosis of sarcoidosis, which is a lung disease with variable effects on pulmonary function. Sarcoidosis cannot be well categorized by disease severity based on GOLD criteria, as the physiological limitations are less clearly correlated with predicted FEV1%, as with COPD. However, this patient was measured through spirometry for FEV1, indicating GOLD3, which was correctly classified in Session 1. He performed three 6MWT sessions at different visits, whereby the second session was 1 month later than first session and third session was 5 months later than first session, beyond the 3-month rehabilitation period. In Session 2, there was mixed classification with 2/3 intervals indicating GOLD2 and 1/3 intervals indicating GOLD3. The patient was likely within transition between status levels, because of rehabilitation progress.

During much later Session 3, there was a clinical note stating that the patient had suffered a significant illness, the reason for return to rehabilitation. This session was the only true mixture, with some classifications of GOLD1/2/3, although 86% of the intervals indicated GOLD3 once again. So lung physiological function likely worsened and the patient regressed back to original status level. Our model would have predicted this correctly, with threshold of 85%. Thus, there is clinical evidence that our model can predict changes in pulmonary function.

PHASE III: MODEL APPLICATION WITH PASSIVE MONITOR DURING DAILY ACTIVITIES

Measuring motion is a potential solution for passive monitoring, which has compliance advantages over active phone applications. That is, patients simply use their personal phones as usual during daily activities, with no special actions necessary. The home trials are more complex than the hospital trials because patients do many different activities throughout the day, while periods when they are carrying their phones are recorded by continuous monitor.

The passive monitor phone app in the home is a reimplementation of the active monitor in the clinic, with a series of cascading filters as input to the predictive model. Such filters require software for activity recognition, beyond simply detecting when the phone is moving. This is a well-studied problem in computer science, using motion sensors to detect which activity a person is performing.[15,21] Moreover, the sensor positions on human body can be distinguished with proper training.[22] Retrieving good walking passively requires two binary filters. The activity filter needs only detect good walking, when the body motion is similar to the walk test. The position filter needs only detect good position, which is when the phone sensor works on a position equivalent to L3. Equivalent positions include empirically pant pockets, coat pockets, belts, and packs, as normal positions for carrying phones.[9] Our model needs walking sessions of 2 min to predict effectively.[23] As the amount is small, the filter can be tuned up for high precision only during periods with definite measurement of good walking on good position.

We have results from several patients who carried a Galaxy S5 mini loaded with our passive monitor for 1 week. The good walking filters on the phone and the server extracted the periods when qualified sessions were taking place. These home

sessions with new patients were utilized as input to the universal model trained by the clinic experiments for status prediction. The female Patient 30 with COPD/CHF and male Patient 35 with COPD were correctly assessed as GOLD3 and GOLD2 during multiple sessions (7 and 75) recognized as good walking, during their daily activities. In addition, we have a prospective patient, who is a GOLD1 smoker with COPD.

Conclusions

In this study, we expand our previous results[7] utilizing phone sensors into accurate prediction of pulmonary function. This requires utilizing demographic features with better training sets and better statistical models, which analyze characteristic motions in addition to step counting. Every patient now has correct modeling of GOLD level, not only session by session (6 min) but also almost every sample by sample (10-sec).

The trained model is providing perfect computation of pulmonary function category (classifying GOLD stage). This is especially true for the first session with pulmonary patients, whereby spirometry has just measured the function so the pulmonary function value can be directly compared with the phone motion value. For senior patients undergoing pulmonary rehabilitation, respiration and motion are closely correlated. This is why the classification is so accurate, in the hospital setting with monitored walk test. The result is encouraging, even when relaxing constraints on our measurements, moving closer to daily activities in the real world, for example, patients bringing smartphones home from rehabilitation, predicting pulmonary functions during daily activities.

In this prototype, we did not expect all population variation could be caught up by a model trained with only 25 patients. However, the model perfectly captured status on 10 of the same patients, with subsequent walk test data in Phase II experiments, including clinical changes. Moreover, the three subjects in Phase III experiments did not participate in model training, yet their status is correctly predicted, even in the more complex passive monitoring.

We are optimistic that full-scale clinical trials are possible to predict health status for cardiopulmonary patients. They need only carry their own phones and do some limited good walking during the day while the passive monitor records unobtrusively. We are actively planning Phase IV for field trials with large populations, extending from 10 patients to 10,000 patients. Care routing can be supported, whereby daily prediction of pulmonary function is utilized to adjust treatment in response to dynamic status, along lines of Phase II and Phase III. More training with more variation will be necessary for the universal model to still be accurate for all patients across all statuses. As the cheapest smartphones are adequate for recording, large-scale population measurement can become an everyday reality for health systems.

Acknowledgments

At the University of Illinois at Urbana-Champaign, the Institute for Genomic Biology provided facilities for software development and data analysis. At NorthShore University HealthSystem, the Department of Respiratory Therapy at Evanston Hospital carried out patient testing and phone recording. All clinical experiments were protected by NorthShore IRB EH15-025 entitled "Mobile phone software MoveSense versus traditional exercise testing to assess health status," with extensions for passive monitors. We thank Brian Edwards and Catherine Zhu for managing and anonymizing phone data and medical records using REDCap archiving system.

Disclosure Statement

No competing financial interests exist.

REFERENCES

1. Schatz B, Berlin R. *Healthcare infrastructure: Health systems for individuals and populations. Series in health informatics*. London: Springer Verlag **2011**.
2. Fox S, Duggan M. Tracking for health. Pew Research Center: Internet, Science & Tech, 2014. www.pewinternet.org/2013/01/28/tracking-for-health/ (last accessed June 26, 2014).
3. Jehn M, Schmidt-Trucksäess A, Schuster T, et al. Accelerometer-based quantification of 6-minute walk test performance in patients with chronic heart failure: Applicability in telemedicine. *J Card Fail* **2009;**15:334–340.
4. Annegarn J, Spruit M, Savelberg H, et al. Differences in walking pattern during 6-min walk test between patients with COPD and healthy subjects. *PLoS One* **2012;**7:e37329.
5. ATS Statement. Guidelines for the six-minute walk test. *Am J Respir Crit Care Med* **2002;**166:111–117.
6. Longo D, Fauci A, Kasper D, Hauser S, Jameson J, Loscalzo J. Part 17: Neurologic disorders. In: *Harrison's principles of internal medicine,* 18th ed. New York: McGraw-Hill, **2011**.
7. Juen J, Cheng Q, Prieto-Centurion V, Krishnan JA, Schatz B. Health monitors for chronic disease by gait analysis with mobile phones. *Telemed J E Health* **2014;**20:1035–1041.
8. Cheng Q, Juen J, Hsu-Lumetta J, Schatz B. Predicting transitions in oxygen saturation using phone sensors. *Telemed J E Health* **2016;**22:132–137.
9. Juen J, Cheng Q, Schatz B. A natural walking monitor for pulmonary patients using simple smart phones. IEEE *J Biomed Health Inform* **2015;**19:1399–1405.
10. Broaddus V, Mason R, Ernst J, et al. *Murray & Nadel's textbook of respiratory medicine*. Philadelphia, PA: Elsevier Health Sciences, **2015**.
11. Rabe K, Hurd S, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* **2007;**176:532–555.
12. Whittle M. *Gait analysis: An introduction*, 4th ed. Woburn, MA: Butterworth-Heinemann, **2014**.
13. Zijlstra W, Hof A. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture* **2003;**18:1–10.

14. Nishiguchi S, Yamada M, Nagai K, et al. Reliability and validity of gait analysis by Android based smartphone. *Telemed J E Health* **2012;**18:292–296.

15. Lu H, Yang J, Liu Z, et al. The Jigsaw continuous sensing engine for mobile phone applications. *ACM SENSYS Int Conf Embedded Networked Sensor Systems.* **2010;**71–84.

16. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* **1995;**20:273–297.

17. Chang C, Lin C. LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol* **2011;**2:27.

18. Keerthi S, Lin C. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Comput* **2003;**15:1667–1689.

19. Hsu C, Lin C. A comparison of methods for multiclass support vector machines. *IEEE Trans Neural Netw* **2002;**13:415–425.

20. Pinto-Plata V, Cote C, Cabral H, et al. The 6-min walk distance: Change over time and value as a predictor of survival in severe COPD. *Eur Respir J* **2004;**23:28–33.

21. Kwapisz J, Weiss G, Moore S. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explor Newsl* **2011;**12:74–82.

22. Park J, Patel A, Curtis D, Teller S, et al. Online pose classification and walking speed estimation using handheld devices. *UBICOMP Int Conf Ubiquitous Computing* **2012;**113–122.

23. Cheng Q, Juen J, Bellam S, et al. Classification models for pulmonary function using motion analysis from phone sensors. *AMIA Amer Med Informatics Assoc annual symposium,* **2016;**401–410.

Address correspondence to:
Bruce Schatz, PhD
Department of Medical Information Science
Institute for Genomic Biology
University of Illinois at Urbana-Champaign
1206 West Gregory
Urbana, IL 61801

E-mail: schatz@illinois.edu

Received: January 11, 2017
Revised: March 2, 2017
Accepted: March 2, 2017
Online Publication Date: March 16, 2017