

Smartphone Derived Movement Profiles to Detect Changes in Health Status in COPD Patients - A Preliminary Investigation

Kelly, D., Seamas, D., & Brian, C. (2015). Smartphone Derived Movement Profiles to Detect Changes in Health Status in COPD Patients - A Preliminary Investigation. In *Unknown Host Publication* IEEE.

Link to publication record in Ulster University Research Portal

Published in:
Unknown Host Publication

Publication Status:
Published (in print/issue): 25/08/2015

Document Version
Author Accepted version

General rights
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.

Download date: 29/03/2024

Smartphone Derived Movement Profiles to Detect Changes in Health Status in COPD Patients - A Preliminary Investigation

Daniel Kelly[1], Seamas Donnelly[2] and Brian Caulfield[3]

Abstract— **Over 3.2 million people in the UK alone have the lung disease Chronic Obstructive Pulmonary Disease. Identifying when COPD patients are at risk of an exacerbation is a major problem and there is a need for smart solutions that provide us with a means of tracking patient health status. Smart-phone sensor technology provides us with an opportunity to automatically monitor patients. With sensors providing the ability to measure aspects of a patients daily life, such a motion, methods to interpret these signals and infer health related information are needed. In this work we aim to investigate the feasibility of utilizing motion sensors, built within smartphones, to measure patient movement and to infer the health related information about the patient. We perform experiments, based on 7 COPD patients using data collected over a 12 week period for each patient, and identify a measure to distinguish between periods when a patient feels well Vs periods when a patient feels unwell.**

I. INTRODUCTION

In the UK, 3.2 million people have the lung disease Chronic Obstructive Pulmonary Disease (COPD), which is the 5th biggest killer in the UK. A key feature of COPD is its progressive nature, with cycles of recurring exacerbations leading to a general decline in patient health status over time. One of the critical challenges in achieving effective management of COPD is being able to identify those patients who are at risk of exacerbation, and implementing intervention strategies as early as possible to prevent or limit the impact of the exacerbation. Current approaches to tracking patient status are based on clinical measures or deployment of questionnaires [1]. Apart from the need for the presence of a healthcare professional, these measures only provide a snapshot measurement of patient status and, in the case of questionnaires, are based on the recall of the patient. At best, they provide a relatively expensive means of episodic measurement of patient health status and, as such, do not facilitate a dynamic model of care where the level of care and the care plan can rapidly be adapted to the patients' need. Therefore, there is a need for development of smart solutions that provide us with a means of tracking patient status, and changes in status over time, without the need for a healthcare professional in the loop at all times. An innovative solution might lie in leveraging data from continuous patient monitoring, in the home and community, to provide a means of tracking patient health status and informing the need for care.

[1]Daniel Kelly is with School of Computing and Intelligent Systems, Ulster University, Northern Ireland dnl.kelly1@gmail.com

[3]Seamas Donnelly is with St. Vincent University Hospital, Dublin Ireland

[3]Brian Caulfield is with Insight Center for Data Analytics, University College Dublin, Ireland

Sensor technology has been used in many different forms in order to perform patient sensing for the purpose of health and well-being monitoring [2]. Studies have mainly involved patients wearing specialized sensors, or specialized sensors being set up in the patients home. While there is a large body of research work in the area of sensors and health and well-being, there exists relatively few works dealing with analysis of patient well-being. While some work does exist in this area, studies have mainly been carried out using very specialized sensor devices under controlled conditions [3]. For example, Salarian et al. [4] used 7 motion sensors, mounted on upper and lower limbs, to develop an instrumented versions of the timed up-and-go test for identifying gait impairments due to Parkinsons Disease. In this paper, we perform a preliminary investigation on the use of smartphones as a method of analyzing COPD patient behaviour and well-being. In particular, we aim to investigate the feasibility of using motion sensor data, captured from a smart phone, to identify links between a patients' activity and a patients' health status.

II. METHOD

A. Data Collection

Patients diagnosed with COPD, who had been admitted to St. Vincents University Hospital Dublin with a COPD related issue during the recruitment phase of this project, were approached to take part in the study. The study commenced for each patient upon discharge from the Hospital. Ethical approval for this study was granted by St. Vincents University Hospital Ethics Committee. Informed consent was provided by all participating patients and patients were given the option to opt out of the study at any time.

A custom Android smart-phone App was developed to record Accelerometer and Gyroscope sensor data. Each patient was given an Android Phone (Samsng Galaxy S3 Mini) with the custom built sensor App installed and a high capacity SD card to store data. No interaction with the phone was required by the patient other than to charge the phone at the end of each day and to put the phone on their person in the morning. Each patient kept the phone with them for a 12 week period. During this time, a research nurse made contact with the patient approximately once per week and recorded health status information as reported by the patient such as visits to the GP, visits to the hospital, changes in prescription and perceived wellness. Upon completion of the 12 week data collection period, a research nurse would visit the patient and retrieve the phone and SD card with the data.

An average of 10.8 GB of motion data was recorded for each subject.

B. Data Analysis

The hypothesis of this work is that smart-phone based motion data, describing a patients daily physical activity, can be linked with the patients health status. The key aim of this work is therefore to carry out preliminary tests to evaluate the feasibility of this hypothesis. Of the 7 patients recruited for this study, 3 reported some type of negative health change over the course of the study while the remaining 4 reported no negative health changes. Of the 3 patients in the negative change group, Patient 1 reported a need for 2 GP visits and a change in medication. Patient 2 reported feeling unwell a few days after being discharged from hospital, required antibiotics and went into respite care for one week. Upon returning home from respite care, the patient reported a brief improvement in perceived wellness before feeling unwell again and returning to respite care for a further 2 weeks. Patient 3 reported that a GP was called to their home 2 times in 2 weeks and resulted in a medication prescription. Patient 3 also report physical pain in the foot. The remaining patients 4,5,6 and 7 reported no GP visit, no hospital visit, no change in medication and no negative perceived wellness.

The key focus of the data analysis carried for this paper was therefore to evaluate the 2 different patient groups and investigate if any significant differences existed the between the negative group of patients and the positive group patients. Furthermore, if any differences exist, analysis would be carried to discover if differences occur during periods when the patients reported a negative health change.

1) Data Processing: In order to describe patient activity, Acceleration data, A_x, A_y,A_z and gyroscope data, G_x, G_y, G_z, was recorded during data collection. During data collection, patients can potentially wear the phone in different positions and ordinations. Due to the unconstrained sensor orientation, useful information such as movement in a particular direction can be lost [5]. In order to overcome this, we utilize the technique described by Kelly et al. [5] where updated accelerometer and gyroscope vectors, $\bar{A}$ and $\bar{G}$, are used, describing motion relative to a global vertical axis.

In order to measure the users activity at a given time t, a sliding window system is used. At time t, a number of different features are calculated based on the accelerometer signals $\bar{A}$, the gyroscope signals $\bar{G}$ and the orientation. We calculate a set of features based on the following measurements:

$\mu(x)$: Mean of signal x. $\sigma(x)$: Variance of signal x - variance is higer for more dynamic activities. $IQR(x)$: Refers to the Interquartile range of signal x. IQR of Gyro can be important for identifying sit to stand activity. $Corr(x, y)$: Refers to the correlation of signals x and y as calculated using the Pearson Correlation. $ROC(x)$: Refers to the rate of change of a signal x.

Using the 8 motion signals and the above 5 measurements, a set of features $f(t)$, comprising 40 features, is used. For every second of data recorded over the course of the 12 week data collection, 5 feature vectors are calculated. All feature vectors for a given day d are then averaged to compute a single feature vector $F(d)$ which describes the average movement profile of a patient for day d.

2) Pattern Analysis: For each patient p, the set of feature vectors $B_p = \{F(1), F(2), ..., F(D_p)\}$ describe the movement of the patient for the duration of the study, where $F(d)$ is the feature vector describing the patients motion on day d and D_p is the total number of days data that was recorded for patient p.

As previously discussed, during the patient data collection, 3 patients reported negative health changes during the study while the remaining 4 reported no negative health changes. Our aim is to discover if there exists any features which are similar within the negative group but are very different to the features which occur in the positive group.

In order to discover aspects of the feature vectors which are common in the negative set of patients but uncommon in the positive set of patients, we implement a machine learning technique which fits this model of feature analysis. If the exact days in which a patient felt unwell were known, it would be possible to train a supervised machine learning classifier on positive and negative data sets. The positive set would store all feature vectors from the exact days each patient was unwell and the negative set would store feature vectors from days when each patient was well. However, an issue with this is that we are required to know the exact days a patient was well and unwell. In this work, we only have a course representation of the patient reported wellness, where patients informed a research nurse of their health information on a weekly basis. More generally, due to the inaccuracy of patient reported wellness, it is not reasonable to assume that an accurate dataset of well Vs. not well days could be collated even if patients were contacted more frequently.

We therefore implement a Multiple Instance Learning (MIL) framework which requires significantly weaker labelling information where labels are not assigned to the individual training instances, but instead assigned to sets of instances named bags. A bag is labelled positive if at least one instance in the bag is positive. Conversely, a bag is labeled negative if all instances in the bag are negative. Maron et al. [6] proposed a Diverse Density MIL method where the goal is to identify a point on the feature space which is similar to features which are common among positive bags but dissimilar to features within negative bags.

In this work, a bag is represented by the set of features B_p for a particular patient p. The bag B_p is labelled positive if patient p reported at least one negative health change. Conversely, a bag B_p is labelled negative if the patient p reported no negative health change. The goal of diverse density is to identify the optimal point on the feature space where all positive bags have common features but where no negative bags have any features. We call this point the target movement profile h.

To identify a target movement profile, a measure of Diverse Density is utilized. The main principal of the Diverse Density framework is made up from probability density mea-

sures $P^+(x = h|B_i^+)$ and $P^-(x = h|B_i^-)$, which compute the density of positive points and the sparsity of negative points, for a given concept movement profile x, respectively. Assuming there exists an optimal target movement profile h, the goal is to identify the target movement profile by simultaneously maximizing the density of positive points and sparsity of negative points over all concept movement profiles x in the feature space.

This is formally described in Equation 1-3, where x is maximized in order to identify a point in the feature space which has a high density of positive points and a low density of negative points.

$$\underset{x}{\operatorname{argmax}} \prod_i P^+(x = h|B_i^+) \prod_i P^-(x = h|B_i^-) \quad (1)$$

$$P^+(x = h|B_i^+) = 1 - \prod_j (1 - P(x = h|B_{ij}^+)) \quad (2)$$

$$P^-(x = h|B_i^-) = \prod_j (1 - P(x = h|B_{ij}^+)) \quad (3)$$

$$P(x = h|B_{ij}) = exp\left[-\sum_k x_s^2(k)(B_{ijk} - x_f(k))^2\right] \quad (4)$$

The individual density probability, $P^+(x = h|B_i^+)$, is modelled on the probability that not all points are different from the concept movement profile. Thus, $P^+(x = h|B_i^+)$ is high if at at least one instance in the bag is close to x. Conversely, the sparsity probability, $P^-(x = h|B_i^-)$, is modelled on the probability that all points are different from the concept movement profile. If every positive bag has an instance close to x and no negative bags are close to x, x will have a high Diverse Density.

The probability that an individual movement profile, B_{ij}, is the same as a concept movement profile is based on a distance between them. Different features will have different levels of importance in terms of accurately measuring the similarity of two movement profiles. The similarity between a feature vector, B_{ij}, and a concept movement profile x is therefore defined in Equation 4 as a weighted distance between individual movement features. Where B_{ijk} is the k'th feature of the j'th feature vector in the i'th bag. The target movement profile, h, comprises a target feature vector, h_f, and a scaling vector component, h_s where $h_s(k)$ is a weighting for the k'th feature.

Since the target movement profile is made up of both a feature vector component and a scaling component, the goal of maximization is to find a combined optimal point in the feature space and an optimal weighting for each individual feature dimension. In this work we use the L-BFGS [**?**] method of gradient descent and optimization in order to identify the optimal target movement profile h which produces the maximum diverse density.

III. Results

As part of the preliminary investigations discussed in this paper, the data sets for the 7 patients were grouped into positive and negative bags. The Diverse Density MIL

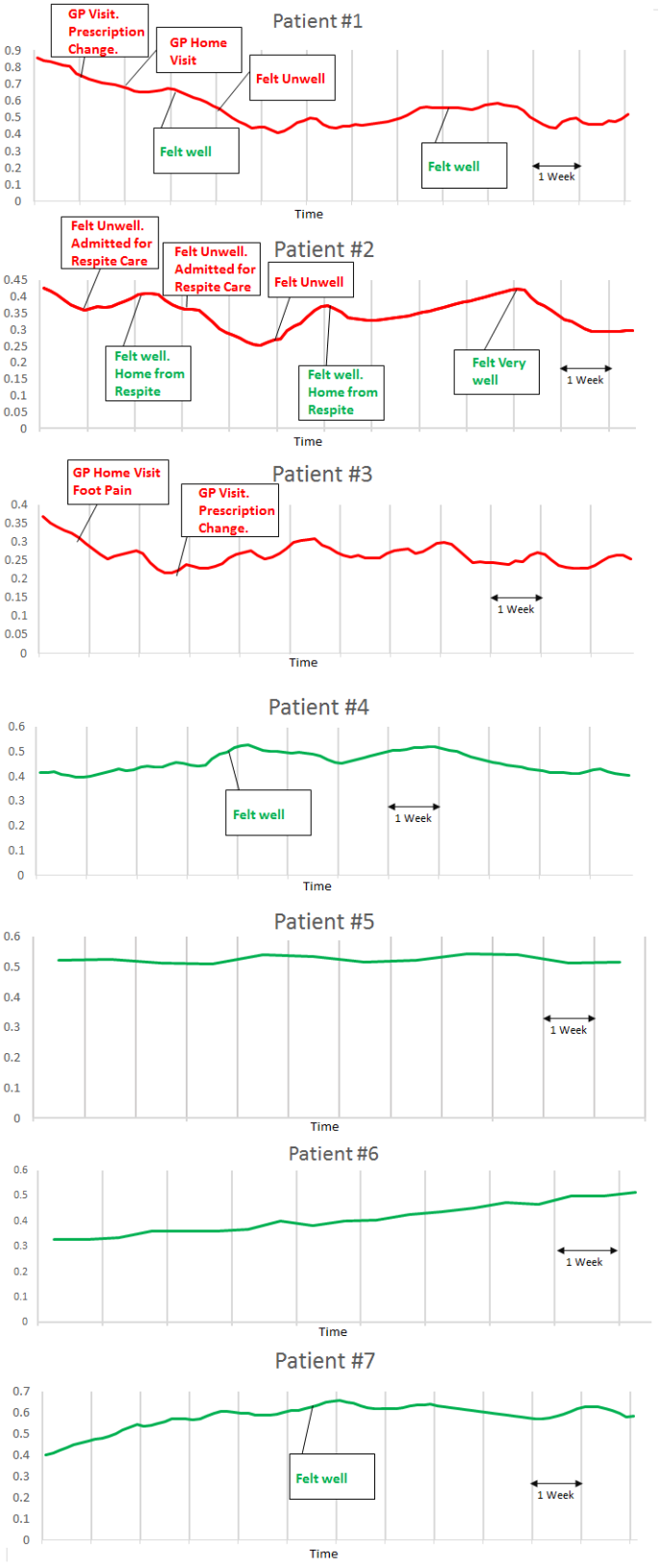


Fig. 1. Target Movement Profile Distances for all Patients

algorithm was then run on the positive and negative bags with the aim of identifying a target movement profile h representing common movement features which occurred among patients who reported negative health changes, but did not occur within the patients who reported no negative health changes. As part of the target movement profile search, a feature point and a feature weighting is identified for each feature in the target movement profile. The feature weight represents the importance of the individual feature. A feature with a high weighting means that the individual feature was common among positive bags but was uncommon within negative bags. In the case of this work, this would mean that the feature was important for identifying negative health changes. Conversely, a low weight would mean that a feature was not important in identifying negative health changes.

Results for this work showed, after running the MIL algorithm on the 7 patient dataset, that 6 features where found to have a significantly higher weighting than all other features. These features, in order of weighting s, are as follows: 1) Average Vertical Acceleration ($s = 0.22$), 2) Average Angular Velocity around Vertical Axis ($s = 0.19$), 3) Variance of Vertical Acceleration ($s = 0.19$) , 4) Average Change in Direction of Vertical Acceleration ($s = 0.16$), 5) Average Orientation Rate of Change ($s = 0.12$), 6) Correlation of Horizontal and Vertical Acceleration ($s = 0.11$). These results indicate that vertical acceleration is a key differentiator between when a patient is well Vs. unwell. This could be attributed to reduced sit to stand activities being performed, possibly caused by increased lying or sitting. Additionally, Angular Velocity around the Vertical Axis was also identified as important. This feature would be high during fast changes in direction while standing or walking. This could indicate that the movement of a patient feeling unwell is much more gradual that that of a patient who is feeling well. This also evident from the fact that Orientation Rate of Change was another feature which was identified.

The next stage of the analysis was to compute the weighted distance between the target movement profile h and each feature vector $F(d)^p$ where $F(d)$ is the feature vector describing the movement of patient p on day d. This distance was computed for all days for each patient. A qualitative analysis of the distance results was then carried out to discover if any trends existed which linked the target movement profile distance measures to that of negative health changes reported by the patients. Figure 1 shows the distance computed for each patient and the positive and negative health occurrences which were reported to the research nurse by each patient. It can be seen that for all patients who reported negative health changes, there is a decrease in the target movement profile measurement around the same time that a negative health change is reported. Additionally, it can be seen that there is an increase in the target movement profile distance around the same time a positive health occurrence is reported. Another interesting observation to be made from the results is that for all patients who did not report a negative health change, the measurements make an overall gradual increase over the course of the data collection. Taking into consideration that the first day of data collection started when each patient was discharged from hospital, the rise in the graph for the patients who did not report negative health changes could potentially be associated with a general improvement in health.

IV. Conclusion

In this paper we have discussed the need for the development of smart solutions that provide a means of tracking COPD patient status, and changes in status over time, without increasing the workload of healthcare professionals. An important aspect of tracking a patients status, is understanding a patients physical activity and, in particular, how it links with a patients' health status. In this work we have carried out a preliminary investigation on how a patients' physical activity can potentially be linked with their health status.

We recorded motion based data from 7 patients over the course of a 12 week period using a Smart-phone and a custom built App. Data analysis was then performed on the data set in order to discover a target movement profile which commonly occurred in the data sets of patients who reported negative health changes but did not occur in the data sets of patients who did not report negative health changes. We found that features such as Average Vertical Acceleration and Average Angular Velocity around the Vertical Axis, among others, where important in identifying negative health changes.

A qualitative analysis of data was also carried out, using data from the 7 patients, and it was observed that there was a consistent decrease in the distance measure during periods where patients reported negative health occurrences. These result indicate that these types of measures could be used to infer information about the status of a patient. While these results do not conclusively state that this method could be utilized to track a patients' health status, as a preliminary study it does provide promising results which warrant further investigation with a much larger patient cohort.

References

[1] F. Pitta, T. Troosters, V. S. Probst, M. A. Spruit, M. Decramer, and R. Gosselink, "Quantifying physical activity in daily life with questionnaires and motion sensors in COPD." *European Respiratory Journal*, vol. 27, no. 5, pp. 1040–1055, 2006.

[2] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation." *Journal of neuroengineering and rehabilitation*, vol. 9, p. 21, Jan. 2012.

[3] K. Hung, Y. T. Zhang, and B. Tai, "Wearable medical devices for tele-home healthcare." *Conference of the IEEE Engineering in Medicine and Biology Society.*, vol. 7, pp. 5384–7, Jan. 2004.

[4] A. Salarian, F. B. Horak, C. Zampieri, P. Carlson-Kuhta, J. G. Nutt, and K. Aminian, "ITUG, a sensitive and reliable measure of mobility," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, pp. 303–310, 2010.

[5] D. Kelly and B. Caulfield, "An investigation into non-invasive physical activity recognition using smartphones." *Conference of the IEEE Engineering in Medicine and Biology Society.*, vol. 2012, pp. 3340–3343, 2012.

[6] O. Maron and T. Lozano-Pérez, "A framework for multiple-instance learning," *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, pp. 570–576, 1998.