AUTOMATIC GENERATION OF PERSONALISED ALERT THRESHOLDS FOR PATIENTS WITH COPD

Carmelo Velardo, Syed Ahmar Shah*, Oliver Gibson, Heather Rutter, Andrew Farmer, and Lionel Tarassenko*

Institute of Biomedical Engineering, University of Oxford, OX3 7DQ, Oxford, UK
Nuffield Department of Primary Care Health Science, University of Oxford, OX1 2ET, Oxford, UK

ABSTRACT

Chronic Obstructive Pulmonary Disease (COPD) is a chronic disease predicted to become the third leading cause of death by 2030. Patients with COPD are at risk of exacerbations in their symptoms, which have an adverse effect on their quality of life and may require emergency hospital admission. Using the results of a pilot study of an m-Health system for COPD self-management and tele-monitoring, we demonstrate a data-driven approach for computing personalised alert thresholds to prioritise patients for clinical review. Univariate and multivariate methodologies are used to analyse and fuse daily symptom scores, heart rate, and oxygen saturation measurements. We discuss the benefits of a multivariate kernel density estimator which improves on univariate approaches.

***Index Terms*—**m-Health, novelty detection, COPD, chronic diseases, digital health

1. INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is a progressive chronic disease that makes it difficult to breathe normally. The main symptoms are cough (usually with mucus discharge), breathlessness, wheezing, chest tightness, and other symptoms such as fatigue and dizziness. The leading cause of COPD is smoking; other factors that can lead to COPD are long-term exposure to air pollution, fumes, dust, and other lung irritants. Patients with COPD are at risk of "exacerbations", a sudden worsening of symptoms which usually continues for several days and can eventually lead to the hospitalisation of the patient. COPD is likely to become the third leading cause of death by 2030 [1].

The effects of COPD, including exacerbations, can be alleviated by improved self-management (for example, training the patient to recognise the onset of an exacerbation [2][3][4]). However, the knowledge transferred to patients during the training sessions provided by national health-care systems is limited [5]. An ideal system to support COPD patients would provide them with informative material and self-management tools adapted to their specific needs.

m-Health can enable the self-monitoring of symptoms and vital signs by patients and provide them with appropriate electronic self-management tools and material. Although self-management is the main objective, there will be occasions when a patient requires intervention by healthcare professionals (HCPs). A second aspect of m-Health systems is therefore the development and use of algorithms to analyse the patient's self-monitoring data and highlight those patients who require review by healthcare professionals. Here the aim is to avoid hospital admissions caused by exacerbations, thus reducing the cost of care and improving the patient's quality of life. In order to avoid overwhelming healthcare professionals with a large number of false alerts, any alerting algorithm should be able to adapt to individual patients and their changing physiology.

1.1 Background - Novelty detection

Novelty detection refers to the analysis of data patterns to identify abnormality based on the knowledge of normal data. Applications of novelty detection include fault detection [6], detection of cancerous masses in mammograms [7] and patient monitoring in high-dependency care [8]. Reviews of the field can be found in [9] and [10], where multiple approaches are presented. Since there is no single optimal approach and performance is highly dependent on the application and its data characteristics, we compared three different methods (two probabilistic methods and one distance-based method) to determine the best approach for analysis of COPD self-monitoring data.

In this paper we firstly describe our robust data collection system for use by COPD patients (Section 2.1) and then introduce the algorithms for analysing the self-monitoring data (Section 2.2). In Section 3 we discuss our results and in Section 4 we present some conclusions.

2 METHODS

In collaboration with the Department of Primary Health Care Sciences, University of Oxford, we developed EDGE (sElf management anD support proGrammE [11]), an m-Health system specifically tailored to patients with COPD. The system design has two aims: to empower patients with tools to support self-management (symptom questionnaire,

* These two authors contributed equally to the work presented

Bluetooth-enabled pulse oximeter, and multi-media content) and to design a scalable framework to monitor patients and generate robust and reliable alerts in the event of a patient's health deteriorating.

The project involved a 6-month pilot phase, during which 18 patients with moderate or severe COPD used our m-Health tablet-based application for self-monitoring and self-management. Table 1 summarises the demographics of the patients in the pilot study. During this study, the Android tablet front-end, the back-end algorithms, and the web front-end were developed and iteratively improved based on the feedback of HCPs and patients. The project is now in its second phase, a randomised control trial involving 165 COPD patients.

Table 1 Demographics of the cohort of COPD patients in the pilot study.

Characteristic	Value
Female/Male	9 / 9
Age *	71 (9)
COPD Severity	6 Moderate 1 Likely severe 9 Severe 2 Very severe
Days in the study *	179.8 (0.3)
Heart rate bpm *	83.2 (18.3)
Blood oxygen saturation (% SpO_2) *	93.5 (4.1)

**Values shown as* mean (std)

2.1 Data collection

We analysed data collected during the pilot study to select the algorithms to use in the subsequent randomised control trial. Each day, at a time of their choice, the 18 pilot study patients used our mobile tablet-based application to complete a symptom diary and record between 30 and 40 seconds of pulse oximetry data. The pulse oximeter (Nonin Onyx II Model 9560) measures both the pulse rate and the peripheral arterial blood oxygen saturation, SpO_2, the values of which are then transferred wirelessly to the Android tablet (Samsung Galaxy Tab2). The symptom diary (see Table 2) was derived from validated COPD management questionnaires and adapted to our use-case. The symptom diary includes subjective questions (e.g. well-being self-assessment, sleep quality, and symptom levels) and objective quantities (e.g. presence of phlegm, and medication intake).

A typical interaction starts with the patient recording his/her symptoms using the diary and then inserting the index finger inside the pulse oximeter probe to measure SpO_2 and heart rate.

Signal quality assessment is of paramount importance in m-Health applications as the measurements are made by the patients at home, not by expert HCPs. Pulse oximeter data is susceptible to movement artefact and incorrect positioning of the sensor on the finger. The application on the Android tablet therefore includes signal quality analysis and advises the patient to remain still while recording heart rate and SpO_2 for 30 seconds. If the recording exhibits artefact at the start, it is extended by 10 seconds; if artefact is still present at the end, the patient is given the option of repeating the measurement.

Table 2 The COPD symptom diary. The answers are mapped to numbers in order to generate an overall score (a higher score indicates worse symptoms).

Question	Range of values
How are you feeling today?	[0, 5]
How is your breathlessness?	[0, 5]
How is your wheeze or chest tightness today?	[0, 5]
Do you have a cough?	yes / no
How is your cough today?	[0, 3]
Are you coughing up sputum?	[0, 4]
What colour is your sputum?	[White, Brownish]
Do you have a cold (such a runny/blocked nose) or sore throat?	yes / no
Did you wake up last night due to breathing problems?	[0, 5]

The 18 patients who successfully completed the pilot study used the system for six months with a high level of compliance, providing answers to the symptom diary questions and using the pulse oximeter on an almost daily basis. In total, 2523 sessions were collected containing symptom diaries and pulse oximetry data. For every diary, a symptom score was computed as the sum of each answer's score (see Table 2). An additional set of questions was used to monitor the medication intake (asking the patient whether they were using their reliever inhaler, and/or taking steroids, antibiotics or a combination of these).

2.2 Algorithms

The thresholds for generating robust and reliable alerts need to be determined. In conventional COPD monitoring systems, the same threshold is applied to all patient data, or an HCP sets a different threshold for each individual patient (often subjectively, without detailed knowledge of the patient's previous data). This can result in a high rate of false alerts, adding to the HCPs' workload and increasing costs. Our data-driven approach aims to provide automatically-generated, patient-specific alerting thresholds which are robust to inter-patient variability.

As soon as patients begin using the COPD software application on their Android tablet, their data are automatically transmitted to a secure server (behind the National Health Service firewall). Once 40 sets of data points (approximately 6 weeks of data) have been collected, personalised alerting thresholds are computed as described below. The 6-week time period was chosen in collaboration with the HCPs in order to allow time for the patient to become

familiar with the system, and for sufficient symptom and vital sign data to be collected to characterise that patient.

The first 40 data points for each of the three variables (symptom score, heart rate and SpO_2) were taken as the training data for both univariate and multivariate analysis algorithms.

2.2.1 Univariate

The univariate algorithm computes the personalised thresholds by estimating the cumulative distribution function (CDF) from the data acquired during the training period. The percentile of interest is then selected from the CDF and used as a threshold to determine whether future data-points are normal or abnormal. Figure 1 summarises graphically the steps of the algorithm. In order to achieve a smooth CDF, the algorithm was implemented in Matlab using the *ksdensity* function. This implementation takes advantage of the work of [12] to compute a smoothed version of the CDF.

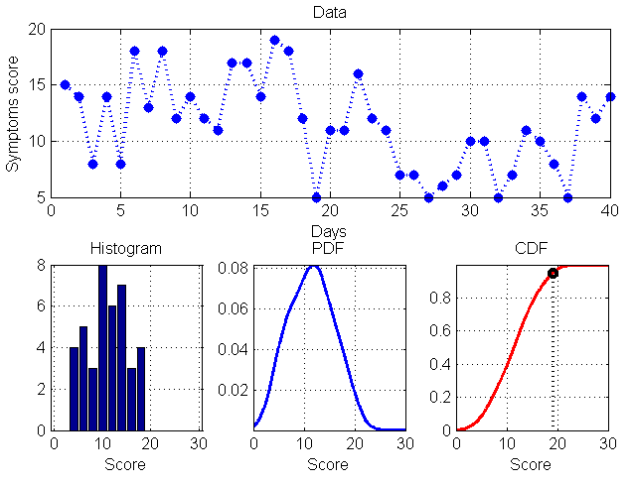

Figure 1 An illustrative method to compute the percentile threshold of a univariate signal. Here applied to the symptom diary data from one pilot-study patient (from the top anti-clockwise): (1) obtain training data points, (2) compute the histogram, (3) estimate the probability density function (PDF), (4) integrate the PDF to obtain the cumulative density function (CDF) and use the 95th percentile as the alerting threshold for symptoms for that patient (score of 19 in this example). Gaussian kernels can replace rectangular ones in order to achieve a smooth PDF and CDF.

2.2.2 Multivariate

The univariate approach assumes the three variables to be independent, which is unlikely to be the case. For example, a decrease in SpO_2 or an increase in pulse rate are likely to be associated with an increased symptom score. In the multivariate approach, we attempt to find a boundary in the 3-dimensional variable space to separate normal and abnormal data optimally.

Distance-based: Each variable is first normalised using the zero-mean unit-variance transform such that $\hat{x} = \frac{x-\mu}{\sigma}$ where μ and σ are the mean and the standard deviation of the variable x in the training data. After normalisation, the Euclidean distance of each data point from the (0, 0, 0) point is computed using $md = \sqrt{diary^2 + pulse^2 + SpO_2^2}$. A threshold is then applied to *md* to determine if the data point is abnormal.

KDE-based: This approach is based on the Parzen windows non-parametric density estimation technique [13]. The model of normality is constructed using an N x 3 dimensional matrix, where N is the number of training data points (symptom score, heart rate and SpO_2). The three variables are again normalised using the zero-mean unit-variance transform. Subsequently, spherical Gaussians are centred on each training data point in the 3-dimensional space and the probability of any data point is computed using equation (1). Since the Gaussian function is smooth, the resulting probability density estimated will also be smooth. In equation (1), σ^2 is a smoothness parameter. It is set to be the mean of local variances, where local variance is estimated by calculating the mean distance to the 10 nearest neighbours [14].

$$p(z) = \frac{1}{n}\sum_{j=0}^{n}\left(\frac{1}{2\pi^{3/2}\sigma^3}\right) e^{-\frac{\left|x-x_j\right|^2}{2\sigma^2}} \qquad (1)$$

$$mp = -\ln(p(z)) \qquad (2)$$

A higher value of $p(z)$ means that z lies close to the distribution of data points from the normal group (in the training set). A novelty score mp is then calculated according to equation (2) which ensures that the lower the value of $p(z)$, the higher will the novelty score be [8].

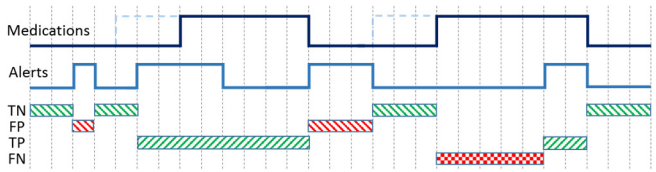

Figure 2 Illustration of how the true/false positives and true/false negatives are defined according to the patient's self-reported medication use.

2.3 Performance Evaluation

In order to validate our alerting algorithms, we treated the self-reported use of medications as an indicator of adverse events (exacerbations). In the context of the current work, an event is identified whenever a patient takes any combination of the three medications (shown by the step change of the blue solid line in Figure 2). For every medication event, there is a 3-day premonitory period (marked by the blue broken line). The 3-day period was deemed by the HCPs as an appropriate

time during which a potential exacerbation could be detected. If no alert is generated when the patient is not taking any medication, we mark those days as true negatives (TN).

If an alert is generated during the premonitory period, or the medication period, then all the days from the day of the alert to the end of the medication event are marked as true positives (TP). If an alert is generated during a medication event, then all days from the day of the alert to the end of the medication event are marked as TP, while the days during the medication event prior to the alert are marked as false negatives (FN). In this case, the days in the premonitory period are marked as TN. In addition, if an alert is generated when the patient is not taking any medication, it is marked as an FN.

For every method, we computed the total number of TP, TN, FP and FN occurrences. Subsequently, for each threshold, we computed the true positive rate (TPR or sensitivity) and false positive rate (FPR or 1-specificity). Finally, the performance of the various methods was compared by evaluating the receiver operating characteristic (ROC) curve and the corresponding area under the curve (AUC).

3 RESULTS AND DISCUSSION

Two of the 18 patients were taking a combination of their medications for over 60 percent (62% and 100%) of the time they spent in the study, and are therefore excluded from this analysis. The remaining 16 patients contributed to 2301 recordings, divided into 640 training points and 1661 testing data points. Of the 1661 testing data points, the patients took medications 97 times (6%), clustered into 15 events.

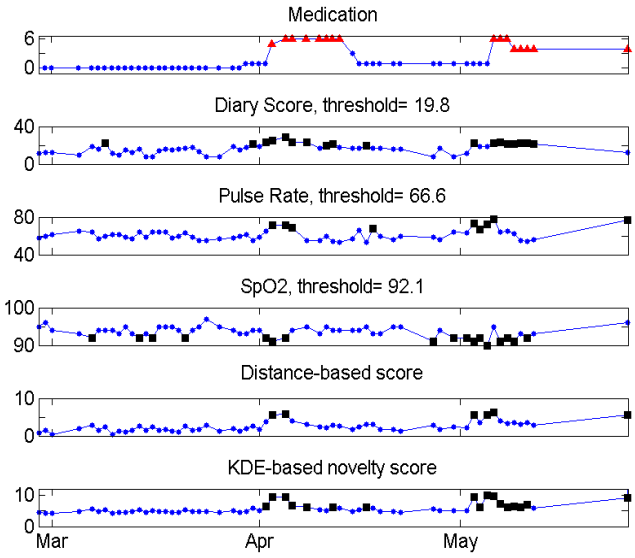

Figure 3 Alerts generated (black squares) using the univariate and multivariate algorithms for one COPD patient in the pilot study. The medication event (during which the patient takes a combination of steroids, antibiotics and or uses the reliever inhaler) is also shown in the figure (medication score ≥ 4).

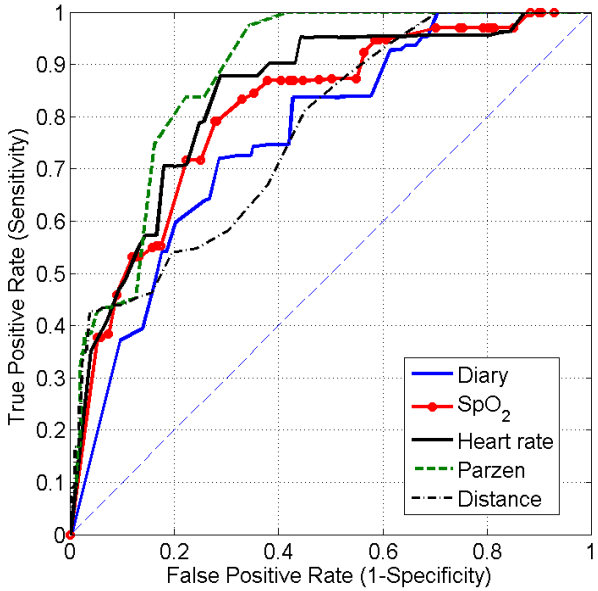

Figure 4 The ROC curves for all the methods described.

Figure 3 shows an example of the alerts generated by each method for one of the patients in the pilot study. The medication events are also shown in the figure using red triangles (medication score ≥ 4 refers to those sessions when the patient took any combination of reliever, antibiotics, and steroids). It can be seen from the figure that the alerts from all the methods are concentrated around the time of the two medication events, and that the KDE-based novelty score may have the best correlation with the medication event.

Quantitative results for all the alerting methods are summarised in Figure 4 and Table 3. The former shows the ROC curves for each of the methods, while the latter gives the area under the curve (AUC) for each method. The KDE-based method outperforms the other methods, followed closely by the Heart Rate and SpO_2 univariate methods. This suggests that heart rate and oxygen saturation are indeed valuable parameters to identify patient exacerbations.

Table 3 The area under the ROC curve for each method employed

Method	Area under the curve
KDE-based (Parzen)*	*0.88*
Pulse	0.84
SpO_2	0.81
Distance-based*	0.77
Symptom score	0.76

*multivariate method

Although the univariate methods were used in the pilot study as the basis for the initial alerting system, we plan to employ the KDE-based method (Parzen windows) in the on-

going randomised control trial because it presents the best trade-off between TPR and FPR.

Since it is not feasible to obtain an objective daily measure of lung infection, we have used the patients' self-reported medication intake as the best available indicator of exacerbation. It is possible that some of the alerts labelled as false positives in fact represent real exacerbations which the patient was not able to identify. Future work could incorporate information on interaction with healthcare professionals (such as phone calls, GP visit or hospital admissions) to help determine more accurately the occurrence and severity of exacerbations.

4 CONCLUSION

We have evaluated methods for automatically setting patient-specific alerting thresholds using a three-dimensional set of data collected with an m-Health application for self-management and remote monitoring of COPD patients. Data collected during a six-month pilot study were evaluated using multiple approaches to identify the best strategy to be used in a subsequent randomised control trial. Data-driven univariate methods already offer advantages over manually-set alerting thresholds, but a KDE-based multivariate novelty detection approach gave the best results in a retrospective analysis of pilot-study data. Future work will include the use of breathing rate (estimated from photoplethysmographic waveform acquired by the pulse oximeter) as an additional parameter in our proposed multi-variate model. We will also adapt our method to identify how often thresholds should be updated to take account of changing physiology over time.

ACKNOWLEDGEMENTS

This publication presents independent research supported by the Health Innovation Challenge Fund (HICF-1010-032), a parallel funding partnership between the Department of Health and Wellcome Trust. The views expressed in this publication are those of the author(s) and not necessarily those of the Department of Health or Wellcome Trust. Additionally, the authors would like to acknowledge the COPD clinical team, and the patients who participated in the study.

REFERENCES

[1] WHO, World Health Statistics 2008, World Health Organization, 2008.

[2] EWMA Bischoff, R Akkermans, J Bourbeau, C van Weel, JH Vercoulen, and TRJ Schermer, "Comprehensive self-management and routine monitoring in chronic obstructive pulmonary disease patients in general practice: randomised controlled trial," BMJ: British Medical Journal, vol. 345, 2012.

[3] JA Wedzicha and J Vestbo, "Can patients with COPD self-manage?" The Lancet, vol. 380, no. 9842, pp. 624–625, 2012.

[4] National Institute for Health and Care Excellence, Chronic obstructive pulmonary disease, NICE, 2010.

[5] R Kessler, E StaŁhl, C Vogelmeier, J Haughney, E Trudeau, CG Lofdahl, and MR Partridge, "Patient understanding, detection, and experience of COPD exacerbations: an observational, interview-based study," CHEST Journal, vol. 130, no. 1, pp. 133–142, 2006.

[6] L Tarassenko, A Nairac, N Townsend, I Buxton, and P Cowley, "Novelty detection for the identification of abnormalities," International Journal of Systems Science, vol. 31, no. 11, pp. 1427–1439, 2000.

[7] L Tarassenko, P Hayton, N Cerneaz, and M Brady, "Novelty detection for the identification of masses in mammograms," in Artificial Neural Networks, Fourth International Conference on. IET, pp. 442–447, 1995.

[8] L Tarassenko, A Hann, and D Young, "Integrated monitoring and analysis for early warning of patient deterioration," British Journal of Anaesthesia, vol. 97, no. 1, pp. 64–68, 2006.

[9] M Markou and S Singh, "Novelty detection: a review part 1: statistical approaches," Signal processing, vol. 83, no. 12, pp. 2481–2497, 2003.

[10] MAF Pimentel, DA Clifton, L Clifton, and L Tarassenko, "A review of novelty detection," Signal Processing, 2014.

[11] A Farmer, C Toms, M Hardinge, V Williams, H Rutter, and L Tarassenko, "Self-management support using an internet-linked tablet computer (the EDGE platform)-based intervention in chronic obstructive pulmonary disease: protocol for the EDGE-COPD randomised controlled trial," BMJ open, vol. 4, no. 1, pp. e004437, 2014.

[12] AW Bowman and A Azzalini. Applied Smoothing Techniques for Data Analysis. New York: Oxford University Press Inc., 1997.

[13] E Parzen "On estimation of a probability density function and mode." Annals of mathematical statistics 33.3, pp. 1065–1076, 1962.

[14] CM Bishop, Pattern recognition and machine learning, vol. 1, Springer New York, 2006.