Online Prediction of Exacerbation in Patients with Chronic Obstructive Pulmonary Disease Using Linear Discriminant Pattern Classification

Morten H. Jensen*, Simon L. Cichosz*, Birthe Dinesen*, Ole K. Hejlesen*

*Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Abstract

Chronic obstructive pulmonary disease (COPD) is a burden on health care because of general care costs. Exacerbations alone cause an additional 23,000 hospitalizations each year in Denmark. Patients seeking treatment for exacerbations often delay consultation for several days after the onset of symptoms. Poor outcomes are often related to failure to seek appropriate treatment therapy. The aim of this study was to investigate whether the use of physiological data is suited for online prediction of COPD exacerbations. Home measurements from 57 patients were analysed and 273 different features were evaluated for their discrimination abilities between periods with and without exacerbations. Results show that if a sensitivity level of 70 % is assessed acceptable the specificity is 95 %, and AUC = 0.73, of the best classifier. Our findings indicate that it is possible to discriminate between periods of exacerbation and periods without. We suggest, that more research in this area should be conducted.

Keywords: COPD, exacerbation, forecasting, rehabilitation.

Introduction

Chronic obstructive pulmonary disease (COPD) is a progressive lung disease in which airways in the lungs are damaged. COPD is characterized by airway obstruction as part of chronic bronchitis or emphysema. It is a major cause of morbidity and mortality in Denmark and around the world. Many people suffer from COPD for years and die prematurely due to the disease itself or its complications [1]. In Denmark, with a population of 5.5 million citizens, COPD has a prevalence of 3-400,000 [2]. The worldwide prevalence of COPD is estimated to 64 million people [3], but the condition may be under-diagnosed [4]. Therefore, the true impact of COPD on morbidity and mortality may be underestimated. COPD is a burden on health care because of general care costs and lost earnings. In addition, exacerbations cause an additional 23,000 hospitalizations each year in Denmark [5] and approx. 20% of all acute admissions [6].

The characteristic symptoms of COPD are chronic and progressive dyspnea, cough, and sputum production. Risk factors for COPD include genetic factors and environmental exposures. The major exposures are tobacco smoke, occupational dusts and chemicals and pollution [7]. Guidelines from the Global Initiative for Chronic Obstructive Lung Disease [7] state that a FEV₁/FVC ratio < 0.7 confirms the diagnosis of COPD. In addition, the precise stage of the disease is classified by spirometer FEV₁:

- I: mild (FEV₁ > 80% predicted)
- II: moderate (50% < FEV₁ < 80% predicted)
- III: severe (30% < FEV₁ < 50% predicted)
- IV: very severe (FEV₁ < 30% predicted or FEV₁ < 50% predicted plus chronic respiratory failure)

Stages III and IV are associated with frequent exacerbations. The median annual frequency of exacerbations using asymptomatic diagnostic system is between 2 and 3 [8]. Studies have shown that a significant proportion of patients experiencing exacerbation fail to fully recover to baseline levels [9] and that poor outcomes are often related to failure to seek appropriate treatment therapy [10]. Patients seeking treatment for exacerbations often delay consultation for several days after the onset of symptoms [11]. Those patients treated early after the onset of symptoms have much better outcomes than those who wait several days [10]. Several studies [12,13,14,15] have tried to predict exacerbation risks on an annual basis for individual patients or patient groups, with partial success. But to our knowledge, no studies have investigated the possibility of predicting COPD exacerbations online, where ‘online’ is defined as prediction on a day-to-day basis. The aim of this study was to investigate whether the use of physiological data is suited for online prediction of COPD exacerbations, in order to start early treatment and improve patient prognosis.

The TELEKAT Project

The data for this study has been obtained from the TELEKAT project (Telehomecare, Chronic Patients and the Integrated Healthcare System). The aim of the TELEKAT project has been to develop and test a preventive tele-rehabilitation program and to enable COPD patients to avoid readmission to hospital by performing self-monitoring and maintaining rehabilitation activities in their own homes. COPD patients with
severe and very severe COPD are included in the study. A telehealth monitor box is installed in the patient’s home for four months. Using wireless technology, the telehealth monitor can collect and transmit data about the patient’s blood pressure, pulse, weight, oxygen level, lung function, etc. to a web-based portal or directly into the patient’s electronic health record. Healthcare professionals and COPD patients can assess the measured data, monitor the patient’s disease and training inputs and provide advice to the patient.

Materials and Methods

Subjects

The subjects are recruited from Aalborg Hospital (Århus University Hospital), healthcare center and general practitioner. The demographic characteristics of the subjects used in this study are as depicted in table 1.

Table 1 - Baseline characteristics of the Tele-rehabilitation groups. The values shown are the mean or median.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tele-rehabilitation group (n=57)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td>Number</td>
<td>23</td>
</tr>
<tr>
<td>Age (years)</td>
<td>69.6 (53.2-82.3)</td>
</tr>
<tr>
<td>FEV1</td>
<td>1.10</td>
</tr>
<tr>
<td>Weight</td>
<td>79.61</td>
</tr>
<tr>
<td>BMI</td>
<td>25.74</td>
</tr>
<tr>
<td>Oxygen saturation</td>
<td>93.33</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>137/79</td>
</tr>
<tr>
<td>Pulse</td>
<td>77</td>
</tr>
<tr>
<td>MRC dyspnoea score</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Due to lack of data density and the scope of online prediction, the parameters of interest are oxygen saturation, pulse and blood pressure. Because exacerbations were not directly registered in TELEKAT, exacerbation dates were indirectly extracted from 1) hospital admissions with SKS code DJ44X (exacerbation, unspecified) or DJ18X (pneumonia, unspecified) and 2) patient journals with indication of administration of either antibiotics or corticosteroids in combination with one or more of the following exacerbation specific symptoms: increasing cough, purulent expectorator or fever. There was a total of 10 exacerbations.

Ethical approval was obtained from the local Ethics Committees (August 27, 2008/ N-20080049). The study was performed according to the Declaration of Helsinki. The project was reported to the Danish Data Protection Agency (granted on August 7, 2008).

Pattern Recognition

Prediction of the exacerbations was managed as a pattern recognition problem. Data from each patient with exacerbation were divided into two groups: 1) a test case, defined as the interval 30 days prior to the exacerbation, followed by a recovery period of 14 days after the exacerbation [17], and 2) control cases in 30-day intervals after the recovery period. The result was 10 test cases and 19 control cases.

Feature Extraction

To classify the case patterns, 273 different features were evaluated for their discrimination abilities. The features are listed in table 2, where par = {systolic blood pressure (sys), diastolic blood pressure (dia), pulse (pul), saturation (sat), saturation plus puls (sat + pul)}, type = {mean (mean), standard deviation (std), skewness (skew), kurtosis (kurt), linear regression (l.reg)} and Di is measurements at time i in days. Notice that only linear regression is calculated for the parameter saturation plus pulse.

Table 2 - Features with ID and formula.

<table>
<thead>
<tr>
<th>Feature ID</th>
<th>Formular</th>
</tr>
</thead>
<tbody>
<tr>
<td>F_par.type[5;0]</td>
<td>type(x){x \in D_i : -5 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[10;0]</td>
<td>type(x){x \in D_i : -10 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[15;0]</td>
<td>type(x){x \in D_i : -15 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[20;0]</td>
<td>type(x){x \in D_i : -20 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[25;0]</td>
<td>type(x){x \in D_i : -25 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[30;0]</td>
<td>type(x){x \in D_i : -30 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[10;5]</td>
<td>type(x){x \in D_i : -10 \leq t &lt; -5}</td>
</tr>
<tr>
<td>F_par.type[15;10]</td>
<td>type(x){x \in D_i : -15 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[20;15]</td>
<td>type(x){x \in D_i : -20 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[25;20]</td>
<td>type(x){x \in D_i : -25 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[30;25]</td>
<td>type(x){x \in D_i : -30 \leq t &lt; 0}</td>
</tr>
<tr>
<td>F_par.type[5;0][10;5]</td>
<td>F_par.type[5;0] \cdot F_par.type[10;5]</td>
</tr>
<tr>
<td>F_par.type[5;0][15;10]</td>
<td>F_par.type[5;0] \cdot F_par.type[15;10]</td>
</tr>
</tbody>
</table>

Since calculation of more of the features requires minimum three data points, this number was set to the minimum data points limit.

The SEPCOR algorithm [18] was used to eliminate features. The algorithm calculates a variability measure for each feature, which is the ability to separate the means of the two classes in relation to the variance of each class. The features are then correlated, and the feature with less variability measure is eliminated if correlation exceeds MAXCOR.
Pattern Recognition Model

The feature values are assumed to be normally distributed since they originate from physiological parameters. However, since the test class only includes 10 cases, the non-parametric method (partly parametric) linear discriminant classification [19] is used. The performance of the classifier is evaluated with the measure of sensitivity and specificity.

To find the optimal feature subset, all combinations of features are tested in different classifiers, and a receiver operating characteristics (ROC) curve is calculated for each classifier. The area under the curve (AUC) is then used as a performance measure for the feature combination.

Results

The subset of features after elimination with the SEPCOR algorithm and MAXCOR = 0.5 is illustrated in figure 1. The illustration is limited to features with \( V(\mathcal{F}) > 0.15 \).

![Variability measures of features after elimination with SEPCOR and MAXCOR = 0.5.](image)

\( F_{sat,reg}\{30;0\} \) and \( F_{sat,std}\{25;0\} \) are the only features where all 10 test cases and 19 control cases are used. For example, for \( F_{sys,kurt}\{20;0\} \), only nine test cases are used, because there are not enough data points (< 3) to calculate kurtosis in the interval of the missing test case. Test of classifiers with different feature combinations is therefore limited to a test of three classifiers, as defined in table 3.

![Figure 2 - ROC curves for the three classifiers.](image)

The ROC curves for the three classifiers are shown in figure 2.

Table 3 - Classifiers necessary to test.

<table>
<thead>
<tr>
<th>Classifier ID</th>
<th>Feature Combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( F_{sat,reg}{30;0} )</td>
</tr>
<tr>
<td>2</td>
<td>( F_{sat,std}{25;0} )</td>
</tr>
<tr>
<td>3</td>
<td>( F_{sat,reg}{30;0}, F_{sat,std}{25;0} )</td>
</tr>
</tbody>
</table>

As can be observed classifier 1 performs worst and classifier 2 best. AUC for classifiers 1 is 0.61, for classifier 2 AUC = 0.78, for classifier 3 AUC = 0.73, which underpins the illustration and confirms that the general classifier performance is best with the use of \( F_{sat,reg}\{30;0\} \) alone.

If a sensitivity level of 70 % is assessed acceptable the specificity is 95 %. That is 7 of 10 exacerbations are predicted, whereas 18 of the 19 control cases are classified as non-exacerbations.

Discussion

The experimental data used in this study were obtained from a project with a different purpose. Because of this, different actors attended to the patients and the sampling of physiological data varied significantly between individuals, resulting in an under-sampling problem. The problem with under-sampling is that physiological noise has a considerable impact.

If, for example, linear regression is calculated from 3 pulse data points and the patient has climbed a stairway immediately before measuring one of the times, this data point may be unnaturally high. If the sampling frequency were higher, the distorting effect of outliers would be reduced and could even be further reduced by low-pass filtering. The missing sampling protocol also results in inequidistant data points, for which reason more potential test cases and control cases are excluded, thereby reducing the significance level. In the sampling process, error measurements were observed from an internal messaging system incorporated in the TELEKAT project for clinical personnel, e.g. when relatives to the patient tried out the system. It is plausible that not all error measurements were registered, which again distorts the results. Finally exacerbations were not directly registered in TELEKAT. In this study, time for exacerbations was deduced from admissions codes and patient journals, a method for finding the test cases assessed as valid. On the other hand, patients might have had exacerbations during the control cases, because all exacerbations were probably not registered. This lack of registration may have caused the classes to overlap.
Linear discriminant functions, which because they sacrifice some performance in order to gain the advantage of simplicity, are often used as initial trial classifiers [19]. If the prediction algorithm were used in practice, the choice of classifier model should therefore be reviewed. The performance reduction, however, is not significant when only one feature is used, since the linear properties do not affect other features. The result with $F_{sat.reg}(5.0)$ is not assessed as unambiguous, because more of the other feature for example the same but in narrower intervals, [5.0], [10.0], [15.0], [20.0] and [25.0] are excluded because the data are too limited. These features could just as well be discriminating, and if more data were available, the feature elimination should be performed again. Furthermore, if new data with higher density were available, features derived from the physiological parameters from the spirometer could probably contain information that could contribute to online prediction.

The prediction algorithm is developed to predict the risk for an exacerbation within the next 30 days. The patient, therefore, does not have any knowledge concerning the onset on a day-to-day basis. As a decision support system, the algorithm can therefore not be used simply to help with self-administration of antibiotics and/or corticosteroids. With training data of higher density, it is maybe possible to identify features that characterize the onset of exacerbations more specifically. If, for example, $F_{sat.reg}(5.0)$ was discriminating, the patient would obtain the risk for the onset of an exacerbation within 5 days.

Our findings indicate that it is possible to discriminate between periods of exacerbation and periods without. Physiological data can therefore be used for online prediction of exacerbation. To optimise the algorithm and thereby develop a system which can support the patient in administering medicine and controlling their disease, it is necessary to identify those features that can characterise the onset more specifically. This feature extraction is not possible with the data used in this study, but it may be possible with data that has a consistent sampling protocol. Such a consistent sampling protocol should include at least one sample per day and preferably equidistant samples. A system capable of predicting the risk within the interval of few days can provide support to the COPD patient in their tele-rehabilitation, helping them to either contact medical assistance in time or to administer their medicine, thereby improving prognosis. Because of these factors, we suggest, that more research in this area should be conducted.

**Acknowledgments**

We wish to thank the COPD patients and relatives participating in the TELEKAT project as well as clinical and industrial partners (for details, see www.Telekat.eu). The TELEKAT project is funded by the Program for User-driven Innovation, the Danish Enterprise and Construction Authority, Center for Healthcare Technology, Aalborg University, and by various clinical and industrial partners in Denmark. Finally, we would like to thank chief physician Carl Nielsen, Aalborg Hospital, Denmark and GP Ove Grann, Vejgaard Healthcare center.

**References**


Address for correspondence
Ole Hejlesen
E: okh@bst.aau.dk
T: +45 2045 9779

Morten Hasselstrøm Jensen
E: mhje06@student.aau.dk
T: +45 2222 6964

Simon Lebech Cichosz
E: +45 2257 7644
T: scicho06@student.aau.dk