ORIGINAL RESEARCH

# Protein data in the identification and stage prediction of bronchopulmonary dysplasia on preterm infants: a machine learning study

Pratheeba Jeyananthan[1] · K. M. D. D. Bandara[1] ·
Y. G. A. Nayanajith[1]

**Abstract** Bronchopulmonary Dysplasia (BPD) is a chronic lung disease mostly affecting the premature newborns who are in the need for oxygen therapy. Main reason for this disease is underdeveloped lungs which need the help of ventilator to expand and breath. This is a very serious disease without a specific test for diagnosis. As treatment is very important for this disease in order to improve the lung function of the baby, on time diagnosis is crucial. Hence, this study checks the potential of protein data in the diagnosis of BPD, and also in the prediction of BPD stage. Mutual information is used in the selection of relevant features of each study. Selected set of features are used with different machine learning algorithms and the accuracies among the models are compared. By this comparison, this study reveals the best number of features in each of the prediction along with the best machine learning algorithm. The highest accuracy value obtained in the diagnosis model shows that it can be used in practice to improve the diagnosis accuracy. However, it seems that prediction of the disease stage is a far complex problem which needs further improvement in its accuracy.

✉ Pratheeba Jeyananthan
pratheeba@eng.jfn.ac.lk

K. M. D. D. Bandara
2018e019@eng.jfn.ac.lk

Y. G. A. Nayanajith
2018e083@eng.jfn.ac.lk

1 Department of Computer Engineering, Faculty of Engineering, University of Jaffna, Jaffna, Sri Lanka

## 1 Introduction

Bronchopulmonary dysplasia (BPD) is the most common and severe respiratory disorder targeting the extremely preterm infants, because of their immature lungs [1, 2]. Even after 50 years of its initial diagnosis, still there are no specific definitions or cures for this disease. It is associated with increased mortality, respiratory morbidity, neurodevelopment impairment and increased healthcare costs [3]. Around 25% of infants with BPD have the risk of BPD-pulmonary hypertension, which leads to high morbidity and mortality [4]. This is a very serious disease affecting the preterm babies who already undergo numerous health complications. At the extreme level, it can lead to the loss of life.

Hence, on time diagnosis and treatment are crucial for this disease. Studies show that the diagnosis of this disease highly depends on the level and duration of oxygen therapy [5]. Compared to initial days, much improvements are there in the understanding of the risk factors of BPD. However, the definition and effective preventive strategies are still challenging and changing time to time. Based on the necessity for respiratory support, this disease can be divided into three stages as mild, moderate and severe [6]. Deciphering the etiopathology of this disease is a critical process including a comprehensive clinical evaluation along with advanced hemodynamic assessment by echocardiography and other cardiac imaging modalities [3, 7].

Even though there are no recommended medication for the prevention of BPD, literature shows that after the identification of this disease, few treatments can be given