



INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv



A Bayesian Approach to Explore Risk Factors for Respiratory Dysfunction in Intensive Care Unit Patient

Norliyana Binti Nor Hisham Shah^a, Normy Norfiza Binti Abdul Razak^{a,*}, Asma Binti Abu Samah^b, Nur Athirah Binti Abdul Razak^a, Agileswari Ramasamy^a, Fatanah M. Suhaimi^c, J. Geoffrey Chase^d

^a Institute of Energy Infrastructure (IEI), College of Engineering, Universiti Tenaga Nasional, Kajang, 43000, Malaysia

^b Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi, 43600, Malaysia

^c Craniofacial and Biomaterial Sciences Cluster, Advanced Medical and Dental Institute, Universiti Sains Malaysia, 11800, Penang, Malaysia

^d Centre of Bioengineering, University of Canterbury, Christchurch, 8140, New Zealand

Corresponding author: *normy@uniten.edu.my

Abstract— Respiratory dysfunction and failure are common in the intensive care unit (ICU); they are often the primary reasons for ICU admission and affect length of stay, mortality, and cost. However, diagnosing respiratory dysfunction requires arterial blood gas values to calculate the partial pressure of arterial oxygen (PaO_2) to a fraction of inspired oxygen (FiO_2) or P/F ratio. These intermittent blood gas values may be difficult to obtain in some patients or where financial resources are limited. Its varying etiologies and lack of other specific biomarkers make diagnosing difficult without this measurement. Thus, in this study, we investigate commonly available parameters in the ICU for the classification of respiratory dysfunction without arterial blood gas values using a Bayesian network, an unsupervised structural learning method. Clinical data from selected patients in the Medical Information Mart for Intensive Care (MIMIC) III v1.4 database ($N > 8900$ patients) is used to create and validate these models. Bayesian network generated using the taboo order algorithm showed a satisfying performance in the classification of respiratory dysfunction. Results are compared to standard diagnosis with P/F ratio. The predictor variables selected could stratify respiratory dysfunction with 80% accuracy and 94% sensitivity. Hence, without using arterial blood gas values, these parameters could identify respiratory dysfunction in 90% of cases using Bayesian networks.

Keywords— Bayesian network; respiratory failure; intensive care unit; machine learning; classification.

Manuscript received 15 Dec. 2022; revised 19 Jun. 2023; accepted 14 Aug. 2023. Date of publication 30 Nov. 2023.
International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



I. INTRODUCTION

Critically ill patients frequently develop organ failure requiring support, such as mechanical ventilation or hemodialysis [1], [2]. Respiratory failure has been identified as the most common organ failure at intensive care unit (ICU) admission, with an associated risk of mortality up to 38.5% [3-5]. It is often characterized by hypoxemia, typically resulting from pneumonia or non-pulmonary sepsis [6], [7].

Factors associated with increased risk of death in respiratory failure include multiple organ failure, oxygenation index, lung injury, cirrhosis, ventilator-associated pneumonia, and high Acute Physiology and Chronic Health Evaluation (APACHE II) score [8]. Many etiologies and the heterogeneous presentation of respiratory failure make it challenging to diagnose and initiate appropriate treatment. For example, respiratory failure can be caused by sepsis,

pulmonary edema, or virology, as with the recent SARS-COV-2 virus [9], [10].

Several biomarkers, such as interleukin-6 (IL-6) and IL-8, have been associated with respiratory failure. However, none have been identified as specific enough for its diagnosis [11]. Normal respiratory function is usually characterized by partial pressures of oxygen (>80 mmHg) and carbon dioxide (<45 mmHg). Acute respiratory distress and failure are defined as P/F ratio less than 300 mmHg where P stands for partial arterial pressure of oxygen (PaO_2), an arterial blood gas value, divided by the fraction of inspired oxygen delivered (FiO_2) [12].

Patients with respiratory failure may deteriorate rapidly; thus, early diagnosis is deemed necessary. In a logistic regression model for respiratory failure patients, mortality prediction requiring mechanical ventilation, age, vasopressor use, platelet count, serum potassium, hemoglobin, highest heart rate and temperature, and PaO_2 were important variables

resulting in area under the receiving operator curve (AUC) of 0.85 when compared to APACHE III [13]. However, these many variables further highlight its diverse etiology and difficult diagnosis.

Thus, machine learning techniques have been applied to predict respiratory failure for clinical decision support systems. Le et al. developed a gradient-boosted tree model to predict the onset of acute respiratory distress syndrome (ARDS) with high accuracy and AUC of 0.905 [14]. Yang et al. investigated four machine learning algorithms for ARDS onset prediction with different subsets of non-invasive parameters, where the highest performance had an AUC of 0.909 with a boosted tree model (XGBoost) using six minimum features. This study also compared different models using optimum features and showed that performance is not always higher with more features [15]. Identifying highly correlated risk factors with respiratory failure using commonly measured variables in the ICU can deliver faster treatment based on past patients' incidences. Hence, there is a need to define the critical and commonly measured clinical variables linked to diagnosing respiratory failure using the standard P/F ratio, which is not often available in real-time as the arterial PaO₂ cannot be continuously measured and thus cannot be used to obtain the earliest diagnosis.

Bayesian network (BN) is a preferred machine learning method, as the interaction between the variables can be viewed explicitly even when the relationships between the variables are unclear. Furthermore, graphical representation can be easily interpreted without machine learning expertise. It also helps with identifying the most significant factors to the node of interest.

The scope of BN approaches in medical diagnosis is fairly documented in breast cancer, acute kidney injury (AKI), and cardiological events. In breast cancer, BN was employed to map specific genes causing cancer metastases [16], while in acute kidney injury, BN could identify the variables that are directly related to AKI [17]. Meanwhile, BN was employed to predict cardiovascular risk using demographics, lifestyle, and laboratory data [18], as well as the risk of mortality in patients undergoing cardiac surgery [19].

BN has been widely used in the pulmonary system, and one such application is investigating the comorbidity of asthma patients. In this study, it was found that asthma, especially in female patients, is highly associated with chronic obstructive pulmonary disease, respiratory failure, hypertension, atherosclerosis, and gastritis [20]. In another application, a comparison was made between BN constructed using expert knowledge and BN elicited by data employing a large database of lung cancer patients. This study used BN to map the pre-treatment variables such as age, number of comorbidities, cancer stage classification, and type of tumor laterality to treatment plans and 1-year mortality. The resulting performance showed that the BN learned using data achieved an AUROC of 0.81 compared to the BN drawn by experts' knowledge with an AUROC of 0.749. This study suggested that surgery has the highest probability of survival in these lung cancer patients [21]. BN has also been used to distinguish between bacterial colds, virus flu, and pneumonia, which share similar symptoms. The probability of diagnosis was updated based on observed symptoms such as cough and

shortness of breath as well as conditions such as going to crowded places and having a chest x-ray [22].

Similarly, BN also investigated the probability of COVID-19 [23] and used it to determine whether or not a patient with respiratory syndrome is due to COVID-19 [24]. For example, a patient's probability of being diagnosed with tuberculosis is higher with symptoms such as fever, cough, and shortness of breath with a chest x-ray image to rule out influenza, COVID-19, and pneumonia [24]. Meanwhile, BN was constructed using risk factors such as age and obesity with conditions including frontline healthcare workers and going to crowded places, together with observable symptoms which are cough, loss of taste and smell, can distinguish the severity of COVID-19 patients into none, mild, and severe using BN [23]. This shows the large scope of BN applications in risk inference, diagnostics, or decision support for medical applications.

In this study, we aim to identify continuously monitored, routine ICU variables and their relationships with respiratory dysfunction and failure defined according to P/F ratio to develop the basis of a real-time diagnostic directly linked to clinically accepted diagnostic definitions.

The main contributions of this study are as follows:

- The utilization of routinely collected variables continuously monitored at the bedside in the intensive care unit does not include laboratory or imaging to classify the occurrence of respiratory dysfunction according to the respiratory component of the SOFA score.
- The graphical representation shows the presumed relationships between the target node and respiratory-related variables.
- The resulting network achieves high sensitivity without the gold standard, which is invasive arterial blood gas values, for the classification of respiratory dysfunction.

II. MATERIALS AND METHODS

A. Related Works

The vast amount and complexity of data in the intensive care unit (ICU) make it possible to apply artificial intelligence (AI) due to the recent surge in computing power and portability. Several AI applications include diagnosis, disease progression prediction, and disease phenotype identification. Within the ICU, much emphasis has been placed on predicting conditions such as hypotension, hypoxia, and respiratory distress, which are usually observed prior to shock. Concurrent with the recent development of COVID-19, AI, and machine learning were utilized to predict respiratory decompensation using clinical data and imaging. Respiratory distress requiring advanced respiratory support, such as mechanical ventilation, increases the risk of mechanical ventilation-associated pneumonia and mortality [25].

In a recent review of acute respiratory failure (ARF) using machine learning, the classification and severity of ARF are defined based on the mechanical support of oxygenation, which can vary from nasal cannula to invasive mechanical ventilation. The review highlighted the prediction of ARF patients requiring invasive mechanical ventilation (IMV), prolonged IMV, and its failure, which resulted in patients' deterioration and mortality. It also highlighted two studies on the prediction of acute respiratory distress syndrome (ARDS),

which is a severe form of ARF. Generally, the machine learning methods used involved logistic regression, neural networks, decision trees, and random forest employing laboratory data and vital signs as predictors. The most common performances reported were on the area under the receiver operating curve (AUROCs), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) ranging from 0.8 to 0.91 [26].

ARDS was also predicted using electronic health records (EHR) data consisting of laboratory and vital signs with a recurrent neural network. In addition to predicting ARDS, this model was used to predict other conditions and outcomes, such as sepsis, hypoxemia, COVID-19, and mortality. The model's performance was measured using AUROC, which was 0.842. This study identified systolic blood pressure, respiratory rate, neutrophils, lymphocytes, and peripheral oxygen saturation (SpO₂) levels of less than 91% as risk factors associated with higher mortality rates [27].

The risk of ARDS within 24 hours in ICU and COVID-19 patients was predicted in ICU and COVID-19 patients using EHR data, including demographics, comorbidities, and laboratory results. An XGBoost machine learning algorithm was able to predict the cases of ARDS requiring advanced respiratory support in these patient cohorts with an AUROC of 0.858 with high sensitivity and specificity. The factors contributing to its high performance were age, SpO₂, blood urea nitrogen (BUN), pH, and respiratory rate [28].

B. Data Collection

This retrospective study used patient data from the Medical Information Mart for Intensive Care (MIMIC) III v1.4 database [29]. MIMIC III is known to be an open-access database consisting of de-identified ICU patient data from the Beth Israel Deaconess Medical Center from 2001 to 2012. The use of this database is under the approval of the Institutional Review Boards of Beth Israel Deaconess Medical Center.

In total, 8702 patients were selected after meeting the following inclusion criteria: age more than 16 years old, at least one measurement of partial pressure of arterial oxygen (PaO₂), P/F ≤ 300 on the first day of ICU admission, and requiring mechanical ventilation during ICU stay. Further variables extracted from this database are patient demographics and typical vital signs. S/F ratio (SpO₂/FiO₂) was computed as a non-invasive measure of oxygenation index. All variables are listed in Table 1. Measured arterial blood gas, PaO₂, was extracted to define the target node of P/F ratio ≤ 300, which is the Berlin definition of ARDS [12] but is not included in the model as a predictor.

TABLE I
DESCRIPTION OF TARGET AND 13 PREDICTOR VARIABLES

No.	Variables	Minimum	Maximum	Mean
1.	Target (PF)	0	1	
2.	Age	16	91	63.64
3.	FiO ₂	20.99	100	55.61
4.	Gender			
5.	HR	31	198	88.95
6.	MAP	1	50	12.25
7.	Diastolic BP	21	296	59.88
8.	Mean BP	26	297	79.68
9.	Systolic BP	31	341	119.86
10.	PEEP	0.1	30	7.62

No.	Variables	Minimum	Maximum	Mean
11.	rr	0.17	68	20.67
12.	SF	51	476.19	191.77
13.	SpO ₂	51	100	97.08
14.	Temp	30.6	42.8	37.1

*HR, heart rate; MAP, mean arterial saturation; BP, blood pressure; PEEP, positive end-expiratory pressure; sf (S/F): SpO₂/FiO₂; temp, temperature.

In total, 13 predictors in Table 1 after the Target (#2-14), including the derived variables, were considered to build BN with the target node (No. 1 in Table 1) defined as PaO₂/FiO₂ (P/F) ratio (binary; P/F > 300 (0), P/F ≤ 300 (1)). In the ICU, the progression of organ failure is usually tracked by the Sequential Organ Failure Assessment (SOFA) score. This score is computed daily for six main organ systems, including respiratory, cardiovascular, renal, hepatic, central nervous system, and coagulation, on a scale of 0 to 4, where 0 represents normal organ function and an increased score shows worsening organ function. The respiratory component of the SOFA score is assessed using the worst daily P/F ratio [30].

The Berlin definition for ARDS, one of the leading causes of respiratory failure among ICU patients, classifies ARDS severity based on P/F ratio. A P/F ratio of less than 300 is considered mild, P/F < 200 is moderate, and P/F < 100 is severe ARDS. Besides the PF ratio, the Berlin definition has added conditions of chest radiograph and PEEP > 5 cmH₂O for ARDS diagnosis [12]. Hence, in this study, P/F < 300 is selected as the threshold for the target node. Before model creation, the predictors are subjected to data discretization into three intervals. This step transforms continuous variables into a specified number of states, minimizing interactions between the nodes so their continuous probability distribution can be established more easily.

C. Bayesian Networks

Bayesian network is a directed acyclic graph (DAG) consisting of nodes and directed arrows [31-33]. The nodes represent the variables, while the arrow represents the influence of one node on another. The direction of the arrow pointing from one node to another shows the parent-child relationship between the nodes [34]. For example, a directed arc from node A to node B can be interpreted as node A being a parent to node B. The graph is acyclic, so following a path from node A will not return to node A where it started. BN is developed with a conditional probability given by equation (1) where we would like to estimate parameter P(A|B) called the posterior probability representing the conditional probability of A given B. P(B|A) is the likelihood of B given A, and P(A) is referred to as the prior probability of A while P(B) is the marginal probability of B, a normalizing constant that can be removed and can be simplified as P(A|B) = P(B|A) * P(A).

$$P(A|B) = P(A) \times \frac{P(B|A)}{P(B)} \quad (1)$$

Figure 1 shows the typical BN. In this network, node C is the child node of both node A and node B. Meanwhile, node D and node E are both children's nodes of node C. The dependencies of a child node depend on the number and states of its parent. Thus, a child node with more than one parent requires the joint probability distribution of its parent. Thus, the conditional probability of node x, P(x) and the parent node

is π while i is the number of parent nodes can be obtained by equation (2)

$$P(x) = \sum_i P(x|\pi_i)P\pi_i \quad (2)$$

The joint probability distribution of all nodes is represented by chain rules given in equation (3).

$$\begin{aligned} &P(x_1, x_2, \dots, x_n) \\ &= P(x_1)P(x_2|x_1) \dots P(x_n|x_1, x_2, \dots, x_{n-1}) \\ &= \prod_{i=1}^n P(x_i | \text{Parents}(x_i)) \end{aligned} \quad (3)$$

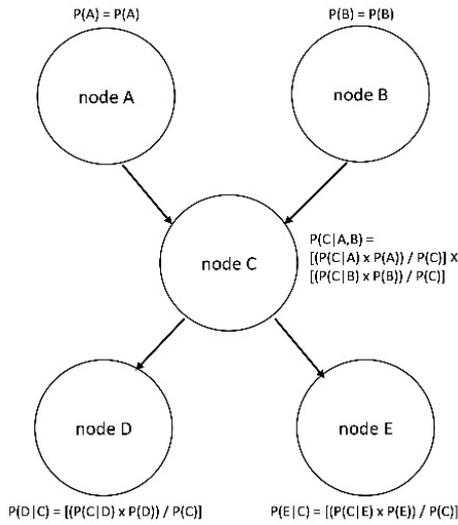


Fig. 1 Typical Bayesian Network

In the unsupervised learning technique for BN creation, where the target node is not specified, the relationships between nodes are built based on posterior probability distribution across the states of the nodes defined by the data discretization technique [35]. Network evaluation was performed on 20% of the randomly selected testing dataset, where the remaining 80% of data was used to create the posterior probabilities and train the network. The trained network is subjected to tenfold cross-validation in which the training dataset was partitioned into ten segments. During cross-validation in the training data, the training was performed on nine segments of the data while holding one segment as a test, and this cycle was iterated ten times. All BNs are built using BayesiaLab version 10.2 (Bayesia S.A.S., France).

The performance of the resulting networks was assessed using precision, reliability, sensitivity, and specificity metrics. The overall precision is the number of correctly identified cases over the total number of cases. The overall precision usually communicates the classification power of the resulting BN. Meanwhile, reliability is described as correctly predicting the number of cases over the number of predictions. Sensitivity is described as the number of positively predicted cases over the number of actual positive cases, while specificity refers to indeed identified negative cases over the actual number of non-respiratory failure cases. As shown in Table 2, True Positive (TP) is correctly identified cases of respiratory failure, whereas False Negative (FN) is the number of respiratory failure cases wrongly identified as not

having respiratory failure. Conversely, True Negative (TN) are positive classification of cases without respiratory failure, while False Positive (FP) is cases without respiratory failure identified as respiratory failure. A higher sensitivity indicates the capability of the resulting network to correctly identify respiratory failure cases, while a higher specificity suggests the potential to distinguish those without respiratory failure from actual cases.

TABLE II
CONFUSION MATRIX STRUCTURE

Predicted	Actual Class	
	Positive (P/F ≤ 300)	Negative (P/F > 300)
Positive (1)	True Positive (TP)	False Positive (FP)
Negative (0)	False Negative (FN)	True Negative (TN)

Data discretization of continuous variables is common in BN, so interactions between discrete states of 195 variables can be built, avoiding complex interactions [35]. Before building the network, equal frequency, k-means, and equal distance data discretization methods were employed [35]. Multiple datasets were developed from different discretization methods, as each technique yields different states of the variables. Equal frequency divides the number of observations into equal numbers across each interval, while equal distance separates the highest number of occurrences with similar traits, often resulting in a higher probability distribution of one interval than the other. Meanwhile, k-means seeks to achieve a normal probability distribution with a centralized one interval comprised of majority distribution, and the remaining number of occurrences is divided equally across other intervals.

Following data discretization, the prepared data were used to build a BN using unsupervised structural learning methods. In unsupervised learning, the network was built intuitively without specifying the target/outcome node [36]. It thus allows visualization of relationships between the variables when the outcome node is not specified, as opposed to relationships concerning the outcome node in the supervised learning method. In this study, three unsupervised learning methods were employed, including the maximum spanning tree (MST), equivalence class (EQ), and taboo order [37]. All networks are evaluated with the specified statistical metrics to assess performance. The construction of BN is shown in Figure 3.

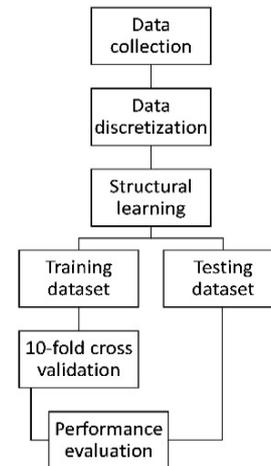


Fig. 2 Model construction for target evaluation

III. RESULTS AND DISCUSSION

As shown in Table 3, the highest mean overall precision result for training and testing data using the equal distance technique was 70.8 %, while equal frequency has the lowest mean overall precision of 41.6%.

TABLE III
OVERALL PRECISION RESULTS OF DIFFERENT DATA DISCRETIZATION METHODS

Overall Precision	Equal Frequency		K-Means		Equal Distance	
	Train	Test	Train	Test	Train	Test
Mean	41.1	41.2	58.9	59.1	75.1	75.2
Minimum	33.5	33.3	42.9	42.8	49.2	49.6
Maximum	76.7	76.6	92.8	92.8	99.7	99.7
On Target Node	76.7	76.6	76.6	76.9	76.7	76.8

Meanwhile, the precision on the target node achieves ~77% across all discretization techniques with slight variation. The k-means and equal distance techniques show SpO₂ has the highest overall precision. The initial probability distribution of discretized data using equal frequency is shown in Figure 3. Continuous data discretized by the equal frequency method is shown because the data for each node is distributed equally based on the number of occurrences. The discrete variables are also shown where respiratory failure (P/F ≤ 300) occurrences were 77%, with one-third of the patients below 58 years old and one-third above 73 years old. The occurrence of respiratory failure as given by the Target node is 76.7%. However, despite the condition to distribute equal numbers of occurrences across specified intervals, 54% have PEEP values less than 5 cmH₂O. This node was re-discretized using the k-means method.

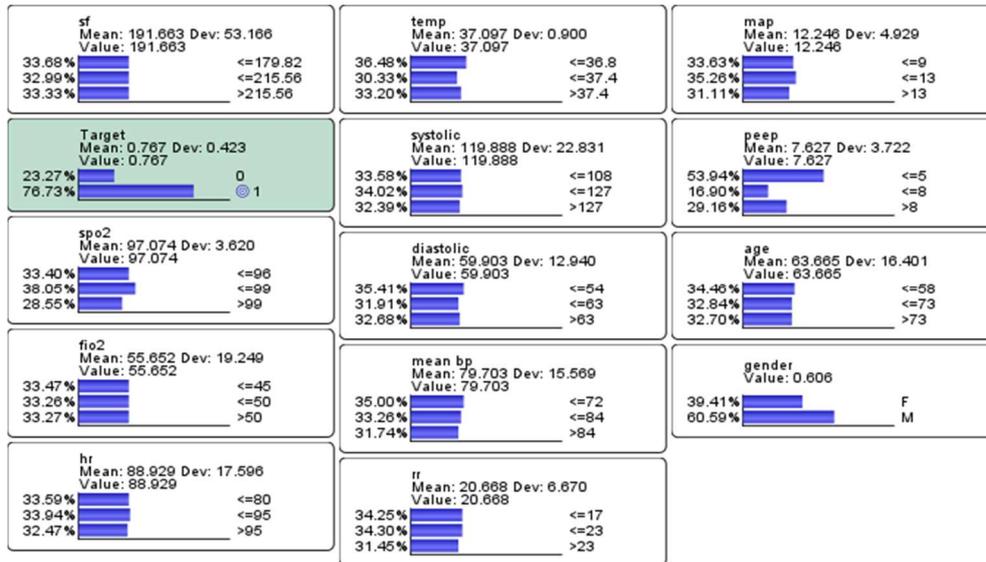


Fig. 3 Initial probability distribution of data discretization using equal frequency

Discretized data were fed into the maximum spanning tree (MST), equivalence class (EQ), and taboo order structural learning algorithms. The performance of these three BNs for each discretization method (nine total outcomes) is shown in Table 4. BN generated using taboo order with data discretized by the equal distance method in the testing dataset has achieved the highest overall precision of 83.5 %. Considering each node in the network as the target allows the computations of minimum and maximum overall precision of the associated

node indicated in brackets in Table 4. BN generated using MST by equal frequency discretization has the minimum overall precision associated with the temperature (temp) node. Meanwhile, using the equal distance method in preparing a dataset for BN generated using EQ shows the maximum precision of a single node was obtained for the FiO₂ node with 99.6%. Generally, a minimum precision of a node of more than 40% suggests they are advantageous for BN creation using unsupervised structural learning methods.

TABLE IV
OVERALL PRECISION RESULTS OF DIFFERENT DATA DISCRETIZATION METHODS

Structural Learning	Training			Testing		
	Equal Frequency	K-Means	Equal Distance	Equal Frequency	K-Means	Equal Distance
MST						
Minimum	41.2	48.3	54.6	40.8 (Temp)	48.4 (HR)	54.1 (Age)
Maximum	99.1	97.7	99.1	99.1 (FiO ₂ , SF)	97.8 (FiO ₂)	99.2 (SpO ₂)
Target	80.1	76.7	76.7	79.5	76.6	76.6
EQ						
Minimum	44.3	50.1	56.2	43.6 (Temp)	50.1 (HR)	55.7 (Age)
Maximum	99.2	99.3	99.6	99.1 (SF)	99.3 (FiO ₂)	99.6 (FiO ₂)
Target	80.5	77.2	77	80.1	77.1	77.1
Taboo Order						
Minimum	45.9	50.5	56.4	45.4 (HR)	50.5 (HR)	56.1 (Age)
Maximum	99.2	99.3	99.6	99.1 (SF)	99.3 (FiO ₂)	99.6 (FiO ₂)
Target	80.5	77.7	77.2	80.2	77.4	77.2

Further examination of the target node in the network demonstrating the highest overall precision on this node was achieved by BNs generated with the taboo order structural learning algorithm employing the dataset discretized using the equal frequency method with 80.5% in the training dataset and 80.2% in the testing dataset, as shown in Figure 4.

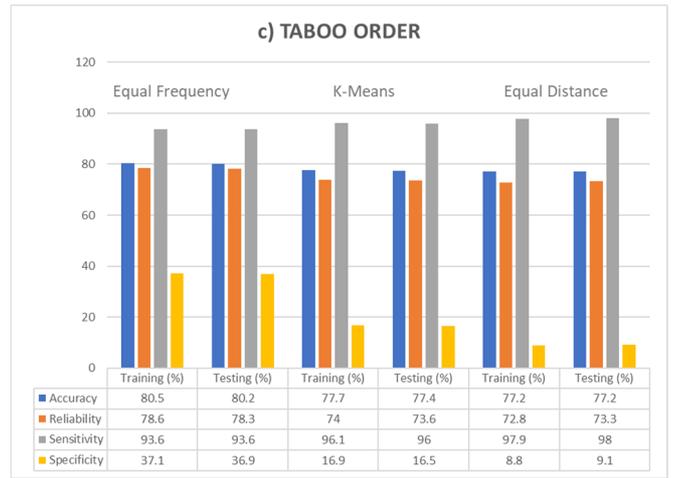
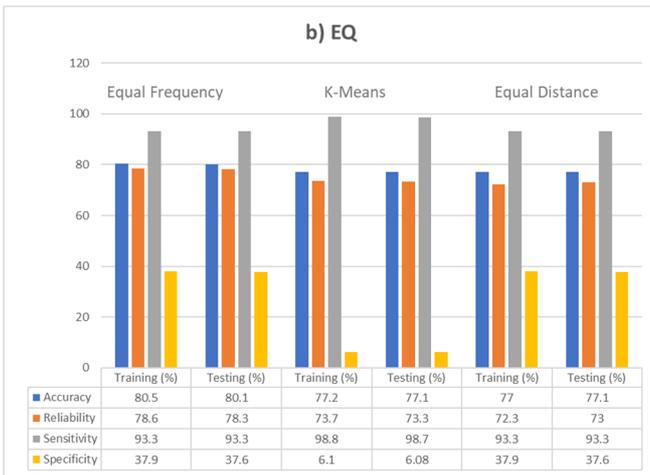
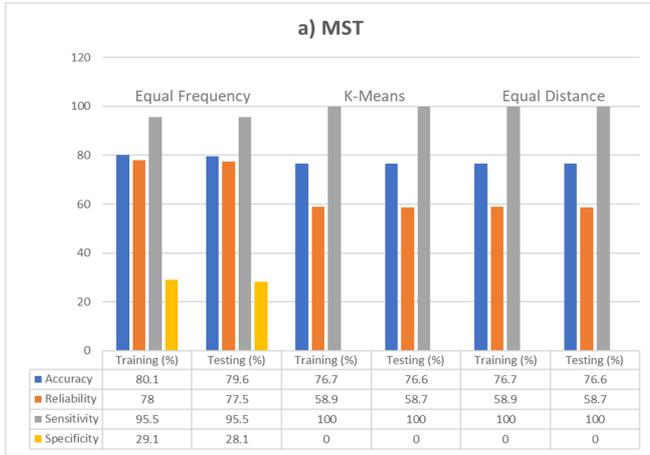


Fig. 4 Target node performance results for a) MST, b) EQ, and c) taboo order structural learning

The highest sensitivity was 100% using datasets generated via k-means and equal distance in BNs generated by MST unsupervised structural learning. This latter result suggests these datasets correctly classified all respiratory failure cases. However, the MST-generated BN with high sensitivity also returns 0% specificity, where all the non-respiratory failure cases were incorrectly classified as respiratory failure. Therefore, taboo order generated BNs using equal frequency discretization are the most suitable prediction model in the absence of PaO₂ with an overall precision of 80% in training and testing data, an accurate prediction (reliability) of 78%, and sensitivity and specificity of 94% and 37%, respectively.

The BN generated using MST consists of 14 nodes and 13 directed arcs. In this BN shown in Figure 5, the classification node denoted as the target is directly connected with SF (S/F ratio) as its parent and SpO₂ as its child node. The SF node stemming from FiO₂ is also parent to PEEP, which connects with rr (respiratory rate) nodes. This MST-generated BN shows that demographic predictors, such as age and gender, are connected away from the target node.



Fig. 5 Bayesian network using maximum spanning tree (MST) algorithm

Meanwhile, the age node junction also connects the relationships between vital sign predictors (HR, bp, Temp). The advantages of this structural learning method are as follows: the relationships between the predictors can be interpreted directly, and visible patterns of variable clustering can be observed to provide further insight.

The BN generated by taboo order unsupervised learning employing equal frequency discretization is shown in Figure

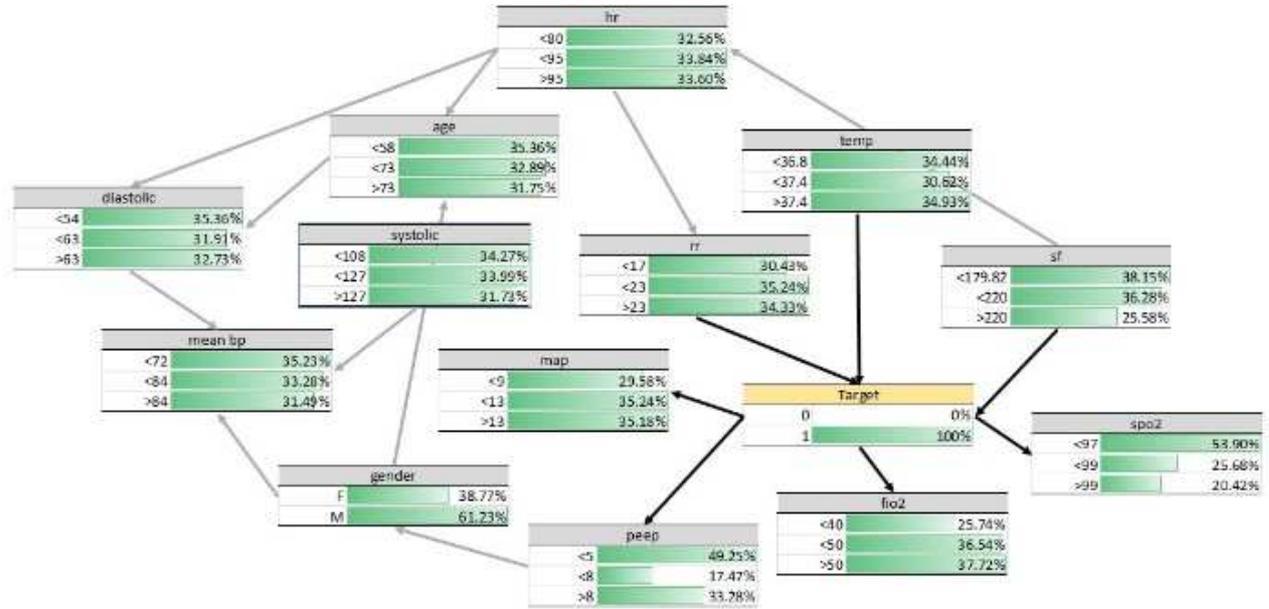


Fig. 6 Bayesian network using taboo order algorithm

Meanwhile, on the other side of the target node, variables relating to blood pressure with demographic variables are grouped at the top. These clusters leave some variables in the middle of the network with the target node, such as heart rate, temperature, MAP, and PEEP, which would vary significantly with respiratory failure. SpO₂ and FiO₂ are both directly connected to the target node, as expected. The derived S/F ratio, which usually corresponds to the P/F ratio, is directly connected to the target as a parent node, similar to the MST-generated BN. The demographic variables, despite not being directly related to the target, offer risk inference for respiratory failure only because they are basic information required for all admitted patients.

The first part of this study compared three different discretization methods with an equal number of intervals for preparing datasets for structural learning. These datasets were employed for unsupervised structural learning using MST, EQ, and taboo order algorithms. The results determined that the taboo order-generated BN using equal distance discretization had the highest overall precision at 83.5%. In contrast, BN generated using an unsupervised learning taboo order algorithm with equal frequency discretization achieved the highest overall precision at the desired target node corresponding to P/F ratio ≤ 300 at 80.5% in the absence of arterial blood gas value, PaO₂, as a predictor. The performances of the specified target node (PF ratio) are consistent with different structural learning algorithms but varied depending on data discretization techniques, suggesting sensitivity to the data discretization technique chosen.

6, with 40 arcs connected among the 14 nodes. Concerning the target node, temp, SF, and rr have parent relationships, while PEEP, FiO₂, SpO₂, and MAP were found to have a child relationship. The color of the nodes can observe clustering patterns within certain types of variables. For example, the variables most related to respiratory function, such as sf, SpO₂, and FiO₂ are clustered near the Target node.

The built BN models showed the parent-child relationship between S/F ratio and the Target node. The use of S/F ratio as substitute to P/F ratio has been debated over time [38-40]. This relationship suggests that S/F ratio is suitable to predict respiratory failure in the absence of P/F ratio as both train and test model showed high sensitivity. Furthermore, in the MST model which limits one parent for every child node, the SF ratio is parent to Target node, with FiO₂ as its parent node. Meanwhile, in the more heuristic search using taboo Order algorithm reveals that SF is parent to SpO₂, FiO₂, and the Target node. The graphical representation of the BN showed that other than the SF ratio, which is highly associated with the target node as observed in MST-generated network, SpO₂, FiO₂, and PEEP, were found to have a direct influence on the target node in the taboo order-generated BN, suggesting their importance in the classification of P/F ratio. Meanwhile, in the absence of PaO₂ as predictors, the variables selected were able to achieve 80% accuracy in the classification of PF ratio, $PF \leq 300$.

The findings in this study were consistent with Ren et al. [41], wherein a large dataset, SpO₂, FiO₂, and PEEP readings can facilitate the assessment of PF ratio with high accuracy. Meanwhile, several variables have been deemed predictive of ARDS, such as low P/F ratio, oxygen saturation (SpO₂), heart rate, hemoglobin, and albumin level [42]. In one study, older age, use of vasopressors, and platelet count were identified to be important predictors among ICU patients with ARDS [11].

This study is limited to all patients who experienced respiratory failure at any point during their ICU stay. Therefore, respiratory and non-respiratory failure patients

have no clear and specific differences. Hence, we recommend studying two patient cohorts, that is, between respiratory failure and non-respiratory failure patients, to better identify which readily available variables are highly associated with respiratory failure.

IV. CONCLUSION

This study demonstrates the use of the Bayesian network to classify respiratory dysfunction defined by $P/F \leq 300$ using commonly available variables in the ICU. Using unsupervised structural learning methods, the Bayesian network mapped the relationship between the predictors and the classification node. Taboo order-generated Bayesian network was found to have the highest overall precision for the network and the classification node. Additionally, SpO_2 , FiO_2 , PEEP, and MAP were found to be highly associated with PF ratio ≤ 300 . This Bayesian network could distinguish respiratory dysfunction cases 90% of the time correctly.

ACKNOWLEDGMENT

The authors thank the Ministry of Higher Education for research grant FRGS/1/2019/STG05/UNITEN/02/1 and Universiti Tenaga Nasional for BOLD Publication Funding and the opportunity to conduct this research.

REFERENCES

- [1] M. Joannidis *et al.*, "Lung–kidney interactions in critically ill patients: consensus report of the Acute Disease Quality Initiative (ADQI) 21 Workgroup," *Intensive Care Medicine*, vol. 46, pp. 654 – 672, 2020.
- [2] D. H. Livingston and E. A. Deitch, "Multiple organ failure: a common problem in surgical intensive care unit patients," *Annals of Medicine*, vol. 27, pp. 13 – 20, 1995.
- [3] G. Bellani *et al.*, "Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries," *JAMA*, vol. 315, no. 8, pp. 788 – 800, 2016.
- [4] G. D. Rubenfeld and M. S. Herridge, "Epidemiology and outcomes of acute lung injury," *Chest*, vol. 131, no. 2, pp. 554 – 562, 2007.
- [5] J. Villar, J. Blanco, and R. M. Kacmarek, "Current incidence and outcome of the acute respiratory distress syndrome," *Current Opinion in Critical Care*, vol. 22, no. 1, pp. 1 – 6, 2016.
- [6] C. L. Auriemma *et al.*, "Acute respiratory distress syndrome-attributable mortality in critically ill patients with sepsis," *Intensive Care Medicine*, vol. 46, no. 6, pp. 1222 – 1231, 2020.
- [7] M. A. Matthay *et al.*, "Acute respiratory distress syndrome," *Nature reviews Disease primers*, vol. 5, no. 1, pp. 1 – 22, 2019.
- [8] S. Bolourani *et al.*, "A machine learning prediction model of respiratory failure within 48 hours of patient admission for COVID-19: model development and validation," *Journal of Medical Internet Research*, vol. 23, no. (2), e24246, 2021.
- [9] Q. Dai *et al.*, "Risk factors for outcomes of acute respiratory distress syndrome patients: a retrospective study," *Journal of Thoracic Disease*, vol. 11, no. 3, 673, 2019.
- [10] C. Wu *et al.*, "Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China," *JAMA Internal Medicine*, vol. 180, no. 7, pp. 934 – 943, 2020.
- [11] R. Houthoofd *et al.*, "Predictive modelling of survival and length of stay in critically ill patients using sequential organ failure scores," *Artificial Intelligence in Medicine*, vol. 63, no. 3, pp. 191 – 207, 2015.
- [12] V. Ranieri *et al.*, "Acute respiratory distress syndrome," *JAMA*, vol. 307, no. 23, pp. 2526 – 2533, 2012.
- [13] Z. Zhang and H. Ni, "Prediction model for critically ill patients with acute respiratory distress syndrome," *PLoS One*, vol. 10, no. 3, e0120641, 2015.
- [14] S. Le *et al.*, "Supervised machine learning for the early prediction of acute respiratory distress syndrome (ARDS)," *Journal of Critical Care*, vol. 60, pp. 96 – 102, 2020.
- [15] P. Yang *et al.*, "A new method for identifying the acute respiratory distress syndrome disease based on noninvasive physiological parameters," *PLoS One*, vol. 15, no. 2, e0226962, 2020.
- [16] F. K. Ahmad, S. Deris, and N. H. Othman, "Improved Bayesian network structure learning for breast cancer prognosis," *Information Sciences and Computer Engineering*, vol. 1, no. 1, 68 – 76, 2010.
- [17] Y. Li *et al.*, "Application of group LASSO regression based Bayesian networks in risk factors exploration and disease prediction for acute kidney injury in hospitalized patients with hematologic malignancies," *BMC Nephrology*, vol. 21, no. 1, pp. 1 – 11, 2020.
- [18] P. Fuster-Parra *et al.*, "Bayesian network modeling: A case study of an epidemiologic system analysis of cardiovascular risk," *Computer Methods and Programs in Biomedicine*, vol. 126, pp. 128 – 142, 2016.
- [19] K. Mazhar *et al.*, "Bayesian networks identify determinants of outcomes following cardiac surgery in a UK population," *BMC Cardiovascular Disorders*, vol. 23, no. 1, 70, 2023.
- [20] Z. Yong *et al.*, "Bayesian comorbidity network and cost analysis for asthma," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 9, pp. 4714 – 4724, 2022.
- [21] M. B. Sesen *et al.*, "Bayesian networks for clinical decision support in lung cancer care," *PLoS One*, vol. 8, no. 12, e82349, 2013.
- [22] C. Jing *et al.*, "Bayesian network based netica for respiratory diseases," *In IOP Conference Series: Materials Science and Engineering*, vol. 435, no. 1, p012022, 2018.
- [23] N. E. Fenton *et al.*, "Bayesian network model for personalised COVID19 risk assessment and contact tracing," *medRxiv*, Preprint posted online January 5, 2021. DOI: 10.1101/2020.07.15.20154286.
- [24] E. O. Edeye *et al.*, "Applying Bayesian Networks to help Physicians Diagnose Respiratory Diseases in the context of COVID-19 Pandemic," *In 2021 IEEE URUCON*, pp. 368 – 371, 2021.
- [25] J. H. Yoon, M. R. Pinsky, and G. Clermont, "Artificial intelligence in critical care medicine," *Annual Update in Intensive Care and Emergency Medicine*, pp. 353 – 367, 2022.
- [26] A. K. I. Wong *et al.*, "Machine learning methods to predict acute respiratory failure and acute respiratory distress syndrome," *Frontiers in Big Data*, vol. 3, 579774, 2020.
- [27] C. Lam *et al.*, "Multitask learning with recurrent neural networks for acute respiratory distress syndrome prediction using only electronic health record data: model development and validation study," *JMIR Medical Informatics*, vol. 10, no. 6, e36202, 2022.
- [28] A. Allen *et al.*, "Risk assessment of acute respiratory failure requiring advanced respiratory support using machine learning," *Research Square*, Preprint posted online June 8, 2022, DOI: 10.21203/rs.3.rs-1668247/v1.
- [29] A. E. Johnson *et al.*, "MIMIC-III, a freely accessible critical care database," *Scientific Data*, 2016, vol. 3, no. 1, pp. 1 – 9, 2016.
- [30] J. L. Vincent *et al.*, "The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure," *Intensive Care Medicine*, vol. 22, no. 7, pp. 707 – 701, 1996.
- [31] R. Montironi *et al.*, "How to develop and use a Bayesian Belief Network," *Journal of Clinical Pathology*, vol. 49, no. 3, pp. 194, 1996.
- [32] J. Pearl, "From Bayesian networks to causal networks," *Mathematical models for handling partial knowledge in artificial intelligence*, Boston, MA: Springer US, 1995, pp. 157 – 182.
- [33] R. Daly, Q. Shen, and S. Aitken, "Learning Bayesian networks: approaches and issues," *The knowledge engineering review*, vol. 26, no. 2, pp. 99 – 157, 2011.
- [34] K. B. Korb, and A. E. Nicholson, "Bayesian artificial intelligence," Boca Raton, FL: CRC Press, 2010.
- [35] J. Dougherty, R. Kohavi, and M. Sahami, "Supervised and unsupervised discretization of continuous features," *In machine learning proceedings*, pp. 94 – 202, 1995.
- [36] B. E. N. Anis *et al.*, "A Bayesian network-based approach to improve the effectiveness of maintenance actions in semiconductor industry," *In PHM Society European Conference*, vol. 2, no. 1, 2014.
- [37] S. Conrady, and L. Jouffe, "Bayesian networks and BayesiaLab: a practical introduction for researchers," vol. 9, Franklin: Bayesia USA, 2015.
- [38] N. Bilan, A. Dastranji, and A. G. Behbahani, "Comparison of the SPO_2/FiO_2 ratio and the PaO_2/FiO_2 ratio in patients with acute lung injury or acute respiratory distress syndrome," *Journal of Cardiovascular and Thoracic Research*, vol. 7, no. 1, pp. 28 – 31, 2015.
- [39] S. Babu *et al.*, "Association between SpO_2/FiO_2 ratio and PaO_2/FiO_2 ratio in different modes of oxygen supplementation," *Indian Journal of Critical Care Medicine*, vol. 25, no. 9, pp. 1001, 2021.

- [40] A. Kumar *et al.*, "Correlation of the SpO₂/FiO₂ (S/F) ratio and the PaO₂/FiO₂ (P/F) ratio in patients with COVID-19 pneumonia," *Medicina Intensiva*, vol. 46, no. 7, pp. 408, 2022.
- [41] S. Ren *et al.*, "Machine learning based algorithms to impute PaO₂ from SpO₂ values and development of an online calculator," *Scientific Reports*, vol. 12, no. 1, pp. 1 – 10, 2022.
- [42] D. Zeiberg *et al.*, "Machine learning for patient risk stratification for acute respiratory distress syndrome," *PloS One*, vol. 14, no. 3, e0214465, 2019.