Sensors & Transducers, Vol. 249, Issue 2, February 2021, pp. 65-71

Sensors & Transducers

Published by IFSA Publishing, S. L., 2021
http://www.sensorsportal.com

The Novel Deep Learning Algorithm and Computer Aided System for Early Prediction of Delirium

[1, *] Jong-ha Lee and [2] Sangwoo Cho

[1] Dept. of the Biomedical Engineering, Keimyung University, Daegu, South Korea
[2] The Center for Advanced Technical Usability and Technologies, Keimyung University, Daegu, South Korea
[*] E-mail: Segeberg@gmail.com

Received: 20 January 2021 /Accepted: 22 February 2021 /Published: 28 February 2021

Abstract: Delirium is generally known to be reversible but can in fact lead to permanent cognitive dysfunction. Despite this, the early detection and reporting rate of delirium in clinical practice is less than 30%. Although there are multiple tools for the early diagnosis of delirium, they require advanced training and are not widely used, which limits their use in real-world practice. This study aimed to use machine learning to classify delirium patients to assess the ease and accuracy of this technique in clinical practice.

Keywords: Artificial Intelligence, Data Processing, Sensor, Mobile Health, Delirium.

1. Introduction

Delirium, also known as acute confusion, is a psychopathological condition that generally occurs over several hours to a few days. The incidence of delirium is thought to be between 22 %–87 %, and varies depending on the population, with the highest incidence among patients in Intensive Care Units (ICU). Delirium occurs suddenly through complicated and variable mechanisms. Fluctuations in consciousness and orientation are characteristics of delirium, but it can also be accompanied by compromised cognition such as decreased concentration and verbal capacity [1].

Both delirium and dementia present as a similar decline in cognitive function, and when assessed by symptoms only, the two conditions are difficult to differentiate based only on the degree of cognitive decline [2]. Nevertheless, delirium and dementia are two independent conditions with different prognoses, and, therefore, it is important to differentiate between the two. The biggest difference between delirium and dementia is the continuity of symptoms. Symptoms of delirium generally occur acutely within hours to days and are corrected within days once the cause has been established and corrected [3]. Symptoms of delirium also show large fluctuations over the course of a day, while symptoms of dementia manifest over months, and their severity tends to be uniform without significant fluctuation. Delirium is a relatively common condition that occurs in 10 %–15 % of all inpatients, particularly in post-operative or elderly patients. The incidence rate of delirium also increases with sudden changes in the environment, such as hospitalization. Indeed, 7 %–47 % of elderly inpatients develop delirium, particularly immediately after admission. Delirium can serve as a warning for physical illnesses and can have a high mortality rate; hence, it is considered a medical emergency. If the underlying pathology is corrected, symptoms can improve within days, but if the cause is uncertain or difficult to correct, it can lead to long-term manifestation. Delirium is generally known to be reversible but can in fact lead to permanent cognitive dysfunction. Despite this, the early detection and reporting rate of delirium in clinical practice is less than 30 %. Although there are multiple tools for the early diagnosis of delirium, they require advanced training and are not widely used, which limits their use in real-world practice. This study aimed to use machine learning to classify delirium patients to assess the ease and accuracy of this technique in clinical practice [4].

There are multiple tools to assess delirium. In South Korea, a translated version of Gaudreau's delirium assessment tool (2005) is available for nurses to apply easily in clinical settings. This tool is based on the Confusion Rating Scale (CRS), with an added category of psychomotor retardation to assess low-activity delirium, which was a limitation of previously available assessment tools. Gaudreau's delirium assessment tool comprises five categories, including inappropriate behavior, disorientation, illusions/hallucinations, psychomotor retardation, and inappropriate behavior, and provides examples of signs or symptoms that represent each category. Nurses can refer to the said examples to assess the severity of the patient's signs or symptoms and assign a score between 0 and 2. The tool is well utilized because it only takes around 1 min to apply and can be used after minimal training to reduce the time and concentration required for history taking. Based on the Korean version of NU-DESC delirium diagnostic criteria, this tool shows specificity of 97 % and sensitivity of 81 %. However, it has limited capacity to predict delirium in post-operative patients due to its low predictability and lack of accurate risk factors. Another delirium assessment tool, the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU), can be used for patients who are intubated or on machine ventilation and have difficulty communicating. The credibility and feasibility of the CAM-ICU has been shown in multiple studies, with sensitivity and specificity of 77.4 % and 72.4 %, respectively. Although the CAM-ICU is a useful tool for the assessment of delirium, the subject needs to be awake in order to answer the questions, which limits its use in the ICU setting in South Korea where patients are often on mechanical ventilation or sedatives [5-7].

Delirium is characterized by a fluctuation in consciousness and disorientation that is accompanied by cognitive dysfunction such as decreased concentration or verbal acumen and psychiatric symptoms. The incidence rate and healthcare burden for delirium can be significantly reduced with early detection, prevention, and intervention. However, due to lack of knowledge or technology, delirium is often not recognized or confused with other conditions such as dementia. Various modalities such as the brain wave test, computed tomography (CT), magnetic resonance imaging (MRI), and delirium assessment tools are being used to diagnose delirium, but the difficulty of applying these tools in a clinical setting leads to an increased demand for healthcare workforce and an increased cost.

As shown in Fig. 1, the recent Korean study on 256 patients with delirium showed that 101 post-operative patients developed dementia within 6 months of surgery, and that elderly patients with history of delirium had a 9-fold increase in the risk of dementia and an increased risk of permanent cognitive dysfunction. Furthermore, the same study showed that, if prolonged, 40 %–50 % of patients expired within 1 year of the onset of delirium (Asan Medical Center). This study, as a precedent study to diagnose and predict delirium symptoms, used RF, SVM, and Random forest machine learning models to analyze the correlation among risk factors that can distinguish delirium patients early. Through this approach, we ultimately sought to assess the feasibility of diagnosing delirium through machine learning [8-9].

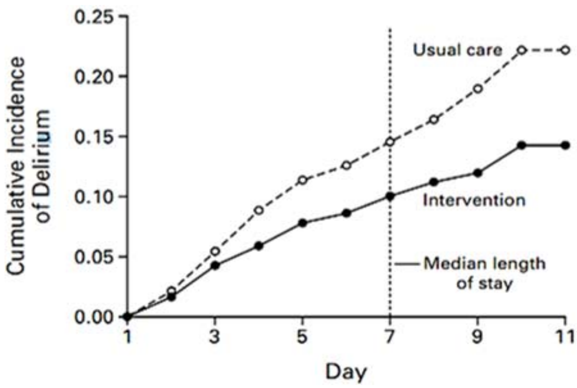

Fig. 1. Cumulative incidence of delirium by group.

2. Materials and Methods

The purpose of this experiment was to gather useful clinical knowledge to assist with the early detection of delirium and to offer an objective capacity index. Prospective cohort study data were used for inpatients at two long-term care facilities in Daegu and Kyeongbuk, with 120 and 100 beds, respectively. The 321 subjects who participated in the 6-month prospective cohort study between October 2016 and March 2017 were categorized into four groups based on the presence of delirium or dementia: patients with no delirium or dementia, those with dementia, those with delirium, and those with delirium leading to dementia. Of the total 321 study subjects, 148 patients who were transferred from a long-term care facility within 12 h of admission, expired during the study period, met the exclusion criteria, or refused to participate were excluded, and the remaining 173 were included in the analysis.

In order to suggest clinical classification criteria for delirium, we re-categorized cohort study patients with no delirium or dementia, or with dementia only to the no delirium group, and those with delirium, or delirium leading to dementia to the delirium group.

Delirium rarely has a single causative factor; rather, it occurs due to a combination of various factors, including age, disease, and environment.

Inouye proposed a multifactorial model that explains the relationship among risk factors for delirium in elderly patients. In this model, the risk factors for delirium were largely divided into predisposing factors and precipitating factors to illustrate the relationship between the two groups. According to this model, patients with multiple predisposing factors such as dementia, comorbidities, and depression showed a high incidence of delirium with only a few precipitating factors such as pain and surgery. Similarly, patients with many precipitating factors showed a high incidence of delirium with only a few predisposing factors. Based on this model, many studies have been conducted on the risk factors for delirium. The major predisposing factors identified from these studies include old age, psychiatric comorbidity, depression, dementia, sleep disorder, nutritional imbalance, alcohol, smoking, drug abuse, and history of delirium. The major precipitating factors identified from these studies include pain, use of restraint, surgery, dehydration, bed sores, and infection.

We aimed to conduct a detailed analysis of the clinical risk factors associated with delirium by analyzing the relationship between each predisposing and precipitating factor.

2.1. Supprot Vector Machine (SVM)

The SVM is a machine learning method that was first proposed by Vapnik (1996) for pattern recognition and data analysis, usually used for classification or regression problems. This model creates a non-probabilistic binary linear classifier model that decides new data categories based on a given data set. The basic principle is to measure the distance between the data from two groups to establish the middle point, and then obtain the optimal hyperplane to learn how to categorize each group. This model defines the decision boundary to classify data and to predict the class of given data by confirming the boundary within which unclassified data fall.

There are countless lines available if data are separated linearly. However, one must determine the optimal line that minimizes the categorization error for data other than the training data set.

In a 3-dimensional space or above, the optimal plane, rather than the line, must be found. In other words, generalizing to the nth dimension, the optimal hyperplane must be found, which in turn becomes the decision boundary. If data cannot be separated linearly, non-linear mapping can be used to convert to a higher dimension to find the optimal linear separation that separates the hyperplane and identify the optimal decision boundary. Mapping in 3-dimension makes it linearly separable. Therefore, using appropriate non-linear mapping in a sufficiently higher dimension will always allow data with two classes to be separated on a hyperplane. The SVM is highly accurate as it can model complicated non-linear decision boundaries, and is also less likely to overfit compared to other models.

2.2. Random Forest (RF)

RF is an ensemble learning model developed by Breiman (2001) that uses a bootstrap method to combine multiple decision trees. In traditional decision trees, the maximum height of the tree can be set by data pruning, but this does not adequately correct for overfitting. RF forms multiple decision trees and processes new data points through each tree simultaneously. Next, it conducts voting within the results categorized by the trees and adopts the result with the most votes as the final categorization result. A tree comprises a hierarchical group of nodes and edges. Nodes are further categorized into internal nodes and terminal nodes (leaf nodes). Decision trees, true to their name, are used for decision making, which divides the decision process into a hierarchy of simple questions. For simple questions, the user can set the parameter, but, for more complicated questions, the model automatically learns both the tree structure and parameters from the training data set. Some trees formed by RF can be overfitted, but this model has the advantage of preventing variability in prediction by forming many trees. The logistic model allows the dependent variable or result value to remain between 0 and 1, regardless of the independent variable. This is obtained by performing odds to logit conversion.

2.3. Logistic Regression

Logistic regression is a probability model that was first proposed in 1958 by a British statistician, D. R. Cox. Logistic regression is a statistical technique that is used to predict the probability of an event by using a linear combination of independent variables. The purpose of logistic regression is similar to that of general regression analysis, in that it aims to represent the relationship between independent and dependent variables in a specific function that can be used for future prediction models.

However, logistic regression differs from linear regression model. The linear regression model can be seen as a classification method, since the dependent variable is categorical, and the result of data input is categorized. The logistic regression model is often used when the dependent variable addresses a binominal problem. If a problem with more than two categories is to be addressed, multinominal logistic regression or polytomous logistic regression is used, and if there are multiple ordinal categories, ordinal logistic regression is used. Logistic regression analysis is widely used for classification and prediction in many fields, including medicine, telecommunication, and data mining.

2.4. K-fold Cross Validation

K-fold cross validation is a method of model analysis in statistics that uses held-out validation (using part of the data set as a validation set) to evaluate the capacity of a model. The limitation of this method is that the small data set compromises the credibility of capacity evaluation for the test set. If the capacity varies by test set, bias occurs in the model evaluation criteria due to coincidence. In order to correct this problem, K-fold cross validation forces all data to be used in at least one test set. This is more time-consuming than normal methods but allows for more learning data sets by classifying data into training and test sets rather than training, validation, and test sets. As a result, it can improve the accuracy for data sets with a smaller number of data points. This study used 10-fold cross validation by dividing the available data into training and test sets and then by dividing training into 10-folds as shown in Fig. 2.

3. Results

3.1. Comparison of Predisposing Factors Between Delirium and Non-delirium Groups

Table 1 shows the comparison of predisposing factors between delirium and non-delirium groups based on the re-classification of earlier cohort study data. The mean age of all patients was 76.9 years, 81.4 years in the delirium group, and 72.7 years in the non-delirium group; thus, the delirium group showed a higher mean age by 8.7 years. In total, 128 of the patients (74 %) were female, and the proportion of female patients was 78.3 %. The Charlson Comorbidities Index (CCI) was higher by 3.5 points in the delirium group, and cognitive impairment (25 %) and medication use (0.8-fold) were also more common in the delirium group compared to the non-delirium group. There was no significant difference in alcohol or BMI between the two groups. Smoking was more common in the non-delirium group. In addition, 46.2 % of patients were transferred from hospital, followed by home and another care facility, showing similar characteristics to the delirium group.

3.2. Correlation Analysis of Risk Factors

The correlation between risk factors identified in previous studies, including age (≥ 65 years), CCI, surgery, sleep disorder, and history of brain damage, was analyzed. The CCI was highly correlated with age and nutritional deficiency, while age ≥ 65 years showed a high correlation coefficient of ≥ 0.5 with medication and cognitive impairment. Bed sores showed high correlation with nutritional deficiency and diaper use. Other factors had relatively low correlation coefficients of ≤ 0.25. Fig. 3 show the results.

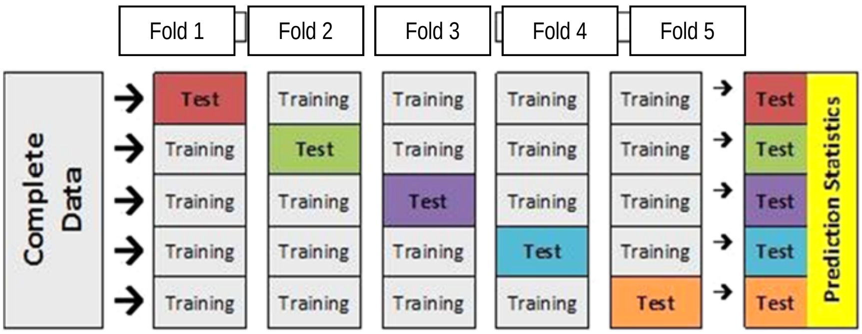

Fig. 2. K-fold cross validation process.

Table 1. Background demographics of participants.

Variable	No delirium	Delirium	Total
Age	72.7 ± 12.1	81.4 ± 8.5	76.9 ± 11.4
Female	63 (70.0)	65 (78.3)	128 (74.0)
BMI (kg/m^2)	21.4 ± 3.4	20.1 ± 3.8	20.8 ± 3.6
CCI	3.1 ± 3.1	6.6 ± 2.9	4.8 ± 3.5
Medication	6.8 ± 3.2	7.6 ± 3.3	7.2 ± 3.2
Smoking	18 (20.0)	11 (13.3)	29 (16.8)
Drinking	8 (8.9)	10 (12.0)	18 (10.4)
Cognitive dysfunction	32 (35.6)	50 (60.2)	82 (47.4)
Home	34 (37.8)	21 (25.3)	55 (31.8)
Hospital	33 (36.7)	47 (56.6)	80 (46.2)
Other facilities	22 (24.4)	15 (18.1)	37 (21.4)
Others	1 (1.1)	-	1 (0.6)

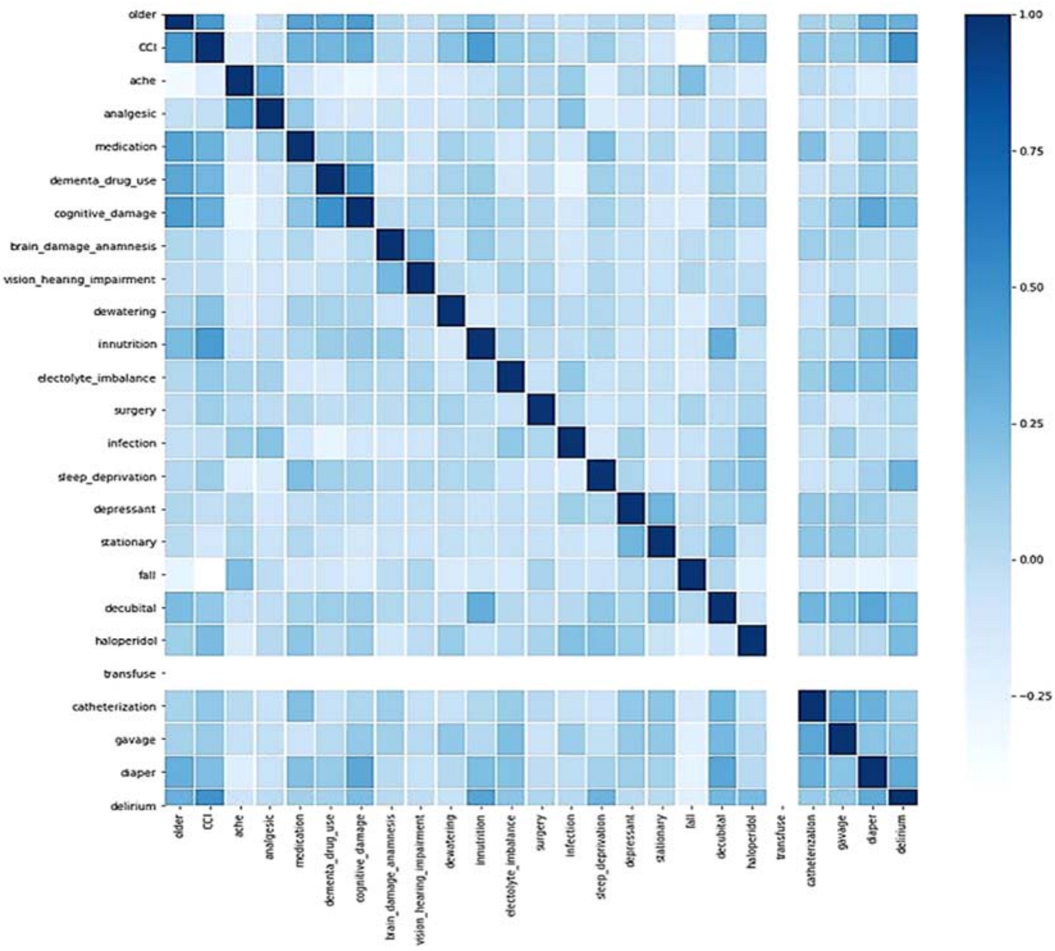

Fig. 3. Correlation analysis of risk factors.

3.3. Risk Factor Analysis of Delirium and Non-delirium Groups

Table 2 shows the results of risk factor analysis comparing the predisposing factors and precipitating factors as selected by previous studies between delirium and non-delirium groups. Predisposing factors that showed higher frequency in the delirium group compared to the non-delirium group included nutritional deficiency, fluid imbalance, sleep disturbance, and delirium medication. Precipitating factors that showed higher frequency in the delirium group included surgery, bed sores, foley catheter, nasogastric feeding, and diaper use. The fall and pain score was greater in the non-delirium group compared to the delirium group. There was no significant difference between the two groups in history of brain damage, infection, dehydration, restraints, etc.

3.4. Model Performance Index Analysis

Fig. 4 and Fig. 5 shows the performance index of each model as assessed by 10-fold cross validation. Of the three models, the RF showed a classification capacity of 0.802 for the delirium group and 0.825 for the non-delirium group. The RF showed the highest performance of the three models, with an average classification capacity as follows: TP rate, 0.813; recall, 0.813; accuracy, 0.813; and ROC area, 0.885. Moreover, the SVM showed an average classification capacity as follows: TP rate, 0.726; recall, 0.724; accuracy, 0.724; and ROC area, 0.724. Logistic regression showed a TP rate of 0.725, recall of 0.724, accuracy of 0.724, and ROC area of 0.791, showing similar results to SVM.

4. Conclusions

This study used machine learning technology to develop a predictive tool for delirium that can be used in a clinical setting to detect and prevent delirium early. Correlation among risk factors categorized as predisposing or precipitating factors in earlier studies was analyzed with group risk factors that were highly correlated. The selected risk factors were than analyzed for their associated with delirium and non-delirium groups to identify clinical factors that can distinguish delirium patients. However, further studies are needed to be able to diagnose delirium based on the identified variables in a real-world clinical setting.

Table 2. Analysis of risk factors for delirium.

Variable	No delirium (n = 90)	Delirium (n = 83)	Total (n = 173)
History of brain damage	8 (8.7)	8 (9.3)	16 (9.1)
Malnutrition	10 (11.9)	37 (46.0)	49 (27.9)
Fluid imbalance	1 (1.3)	9 (8.8)	9 (5.1)
Dehydration	3 (3.1)	3 (3.7)	6 (3.5)
Surgery	3 (3.2)	9 (8.8)	11 (6.0)
Infection	11 (12.2)	12 (13.5)	22 (12.9)
Sleep deprivation	11 (11.8)	33 (37.1)	42 (24.0)
Restraint	1 (1.1)	1 (1.2)	2 (1.2)
Immobilization	3 (3.3)	3 (3.6)	6 (3.5)
Fall	49 (54.4)	26 (31.3)	75 (43.4)
Bedsore	9 (10.1)	26 (29.8)	33 (19.6)
Medication use for delirium	3 (3.1)	16 (18.3)	18 (10.7)
Foley catheter	4 (4.2)	10 (12.2)	14 (8.3)
Nasogastric feeding	4 (4.1)	11 (13.3)	15 (8.7)
Diaper	23 (24.8)	51 (61.2)	74 (42.1)
Pain score	3.2 ± 1.2	2.9 ± 1.3	3.0 ± 1.3

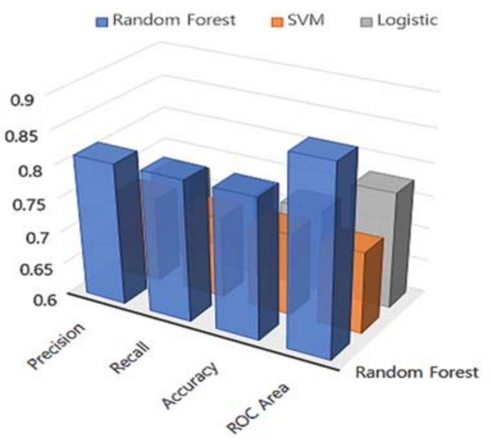

Fig. 4. Performance index analysis.

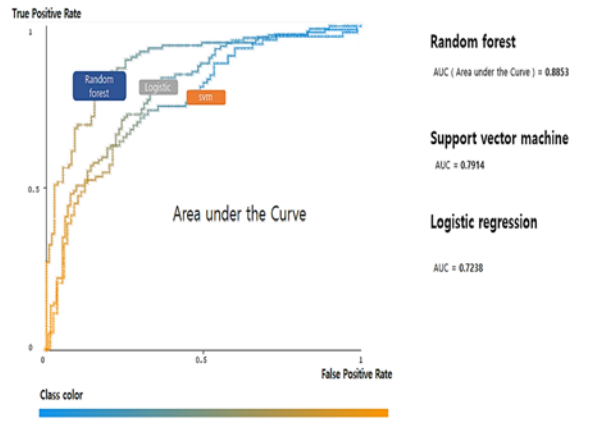

Fig. 5. Comparison of ROC curve.

Three models of machine learning technique that show high performance, RF, SVM, and logistic regression, were used to assess the selection accuracy of machine learning for delirium. A limitation of this study is that the total sample size was small at 173 patients: 83 patients in the delirium group and 90 in the non-delirium group. The inclusion of larger volumes of data would greatly improve the accuracy of the model. Future studies should acquire more data to improve accuracy, and to select classification criteria of the risk factors derived in this study and compare them to other detection tools to assess the clinical applicability of the model. If the capacity of this model is increased, and the accuracy of clinical categorization criteria is improved, healthcare cost and healthcare work force demand for the management and prevention of delirium can be greatly reduced, and treatment efficacy can be improved.

Acknowledgements

This work was supported by the Technology Innovation Program (Development of a novel imaging platform for the screening of cervical cancer; fast speed digital scanner and A.I. engined 3D image reconstruction algorism, B20180275) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea).

References

[1]. Yu Ki Dong, *et al*., Delirium in acute elderly care unit; prevalence, clinical characteristics, risk factors and prognostic significance, *Journal of the Korean Geriatrics Society*, Vol. 9, Issue 3, 2005, pp. 182-189.

[2]. Lee Eun-Joon, *et al*., Risk factors related to delirium development in patients in surgical intensive care unit, *Journal of Korean Critical Care Nursing*, Vol. 3, Issue 2, 2010, pp. 37-48.

[3]. Saczynski J. S., Marcantonio E. R., Quach L., Fong T. G., Gross A., Inouye S. K., Jones R. N., Cognitive trajectories after postoperative delirium, *New England Journal of Medicine*, Vol. 367, Issue 1, 2012, pp. 30-39.

[4]. Maldonado J. R., Pathoetiological model of delirium: a comprehensive understanding of the neurobiology of delirium and an evidence-based approach to prevention and treatment, *Critical Care Clinics*, Vol. 24, Issue 4,

2008, pp. 789-856.

[5]. Lin S. M., Liu C. Y., Wang C. H., Lin H. C., Huang C. D., Huang P. Y., Kuo H. P., The impact of delirium on the survival of mechanically ventilated patients, *Critical Care Medicine*, Vol. 32, Issue 11, 2004, pp. 2254-2259.

[6]. Bergeron N., Dubois M. J., Dumont M., Dial S., Skrobik Y., Intensive Care Delirium Screening Checklist: evaluation of a new screening tool, *Intensive Care Medicine*, Vol. 27, Issue 5, 2001, pp. 859-864.

[7]. Maldonado J. R., Pathoetiological model of delirium: a comprehensive understanding of the neurobiology of delirium and an evidence-based approach to prevention and treatment, *Critical Care Clinics*, Vol. 24, Issue 4, 2008, pp. 789-856.

[8]. Ouimet S., Riker R., Bergeon N., Cossette M., Kavanagh B., Skrobik Y., Subsyndromal delirium in the ICU: evidence for a disease spectrum, *Intensive Care Medicine*, Vol. 33, Issue 6, 2007, pp. 1007-1013.

[9].Breitbart William, Christopher Gibson, Annie Tremblay, The delirium experience: delirium recall and delirium-related distress in hospitalized patients with cancer, their spouses/caregivers, and their nurses, *Psychosomatics*, Vol. 43, Issue 3, 2002, pp. 183-194.