Health Informatics Meets eHealth
D. Hayn and G. Schreier (Eds.)

doi:10.3233/978-1-61499-759-7-32

Development and Validation of a Multivariable Prediction Model for the Occurrence of Delirium in Hospitalized Gerontopsychiatry and Internal Medicine Patients

Diether KRAMER[a,1], Sai VEERANKI[b], Dieter HAYN[b], Franz QUEHENBERGER[c], Werner LEODOLTER[a], Christian JAGSCH[a] and Günter SCHREIER[b]
[a]*Steirische Krankenanstaltengesellschaft m.b.H. (KAGes), Graz, Austria*
[b]*AIT Austrian Institute of Technology, Graz, Austria*
[c]*Medical University of Graz, Austria*

Abstract. Delirium is an acute confusion condition, which is common in elderly and often misdiagnosed in hospitalized patients. Early identification and prevention of delirium could reduce morbidity and mortality rates in those affected and reduce hospitalization costs. We have developed and validated a multivariate prediction model that predicts delirium and gives an early warning to physicians. A large set of patient electronic medical records have been used in developing the models. Classical learning algorithms have been used to develop the models and compared the results. Excellent results were obtained with the feature set and parameter settings attaining accuracy of 84%.

Keywords. Delirium, hospitalized patients, predictive model, electronic medical record

1. Introduction

1.1. Medical Context

Delirium is an acute confusion condition which is a serious neuropsychiatric syndrome [1]. The consequences are adverse in terms of morbidity and increased mortality [2]. Delirium is common in elderly patients [3]. In domestic circumstances, the prevalence of delirium in the elderly population aged above 65 years old is 1-2% and the same increases to 6-56% in the hospital admission settings [3], [4]. Approximately 15-30% of elderly patients are identified with delirium on admission and ~56% will develop delirium during their stay in hospital [4].

Delirium is often misdiagnosed in hospitalized patients [1]-[4]. If delirium could be detected earlier, in 30-40% of the cases, delirium could be avoided by reducing risk factors such as adverse effects of medications, complications from procedures, immobilization, dehydration, poor nutrition and sleep deprivation [3, 5]. Delirium prevention could avoid health care related costs and prolonged hospital stays [6]. Due to

[1] Corresponding Author: Dr. Dieter Kramer, KAGes-Management, Informations- und Prozessmanagement, A-8010 Graz, Billrothgasse 18a, E-Mail: diether.kramer@kages.at

these facts, the occurrence of delirium has been identified as one of the markers for the quality of care and patient safety [6].

1.2. Early detection and prediction of delirium

There are a number of studies on predicting delirium which evaluated different risk-stratification cohort rankings, such as Folstein Mini Mental State Examination (MMSE) scores [7], Clock Drawing Test (CDT) [6], Confusion Assessment Method (CAM) [8], CAM- Intensive Care Unit (CAM-ICU) [8], Delirium Assessment Scale (DAS) [7], etc.

In a literature review, we have identified several relevant papers concerning application of machine learning techniques on delirium prediction that have been published in the last 10 years.

In a recent publication [8], Wassenaar et al. used CAM-ICU to assess the patient delirium conditions and developed a regression model. They achieved an Area Under the Receiver Operating Characteristic (AUROC) of 0.70 with sensitivity and specificity of 62% and 67% respectively in the 0-1 day stay in ICU. These measures increased to AUROC = 0.81 with a sensitivity of 78% and specificity 68% with six days of stay. The study cohort consisted of 2,914 patients in which 1,962 were included in development dataset and 952 were included in validation dataset.

In a similar study [9] with 397 patients who stayed at internal medicine ward, a model has been developed based on a rule derived from CAM. The model achieved an AUROC = 0.85 with sensitivity and specificity of 80% and 90%, respectively.

Several other studies were published on predicting delirium in the years before and a systematic review of risk-stratification models has been published by Newman et.al [10] in which the authors have considered different risk factors and derived rules for predicting delirium [7, 9, 11, 12]. However, current models are not suitable for hospitalized patients in general, the population considered was less in numbers and the accuracy was lower in non-obvious cases.

1.3. Objectives

It was the aim of our study to develop and validate a generalized predictive model irrespective of cohort group but solely on available medical records. Therefore, a large set of records from patients diagnosed with delirium should be analyzed. Delirium induced by alcohol and other psycho active substances should be excluded (i.e. delirium with ICD-10 code F05). The idea was to develop a generalized predictive model to identify the patients who are susceptible to delirium during their hospitalization period, which could alert healthcare practitioners early for keeping a sight on delirium progress.

2. Materials and Methods

2.1. The data set

We used routine data stored within the Hospital Information System (HIS) of the Steiermärkische Krankenanstaltengesellschaft m. b. H (KAGes), a regional health care provider in Styria, one of the nine provinces of Austria. Our dataset consisted of retrospective data of hospitalized patients from gerontopsychiatry and internal medicine departments. This dataset is a part of the HIS of KAGes, which consists of the longitudinal health records from about 90% of the 1.2 million Styrian inhabitants – covering hospital stays and outpatient visits from small standard hospitals to the university hospital in Graz over a period of nearly 15 years.

Inclusion criteria:

- Patients who were diagnosed with ICD-10 Code F05 (delirium due to known psychological condition)
- Time of first documentation of the ICD-10 Code F05 in the period from 01/01/2013 to 30/10/2016
- In-clinic stay at one of the KAGes hospitals

Exclusion criteria:

- For those patients where the day of the diagnosis of delirium occurred within the last two days of a hospital stay we assume that the delirium occurred already earlier and was recorded later. Therefore, we excluded such hospitalizations altogether.
- All data from delirium patients without any patient record before the time of diagnosis – for such patients, no data to be used for prediction (before the diagnosis of delirium) was available

According to the inclusion criteria, we identified approx. 3,000 delirium patients. After applying the exclusion criteria, 2,221 delirium patients remained.

We used the complete record of each of the patients identified according to these criteria for our study, i.e. all data from in-clinic stays and outpatient visits within one of the KAGes hospitals. All the data recorded on the day of the diagnosis of delirium and afterwards were excluded in order to simulate a prospective setting.

2.1.1. Control group

A control group of randomly selected patients without delirium was defined. The control group consisted of 7,000 patients who had at least once been hospitalized at an internal medicine department after 01/01/2013. For minimizing diagnostic bias we limited this group to internal medicine patients only.

2.1.2. Feature set

Data used for modelling consisted of:

- Demographic data (e.g. age, gender, etc.)
- Diagnosis: Several relevant ICD-10 diagnoses based on bivariate proven statistical association were considered, e.g. dementia. Furthermore, we calculated the Charlson Comorbidity Index for each patient from the recorded diagnoses. Suspected diagnoses that have not been confirmed in the KAGes HIS at hospital discharge by the attending physician were not considered.
- Procedures provided to the patient as documented according to the Austrian DRG System were taken into account but we excluded procedures which have been provided too rarely (<10). The selected procedures have been aggregated to their group level.
- Administrative data concerning the patient stay such as patient transfers, admissions and discharges at different departments, number of hospital admissions within the last two years, number of cases (visits to a hospital), longest stay at the hospital, etc.
- Laboratory data (e.g. CRP levels, etc.)

- Nursing assessment (e.g. eating and sleeping disorders, respiratory and communication problems etc.)

2.2. Data extraction / pre-processing

The analysis of the learning models has been done on the data obtained from the KAGes HIS *openMEDOCS*, which is based on IS-H/i.s.h.med information systems, implemented on SAP platforms. Due to the size of the data and the requirements for analysis (both for controlling as well as for medical, quality and efficiency issues) SAP HANA was chosen as the data warehouse platform.

The queries resulted in data set with 8,561 patients and 858 features. Categorical features, such as diagnoses or procedures, were expanded to one boolean feature per possible value. Similarly, ICD-10 codes from main and sub diagnoses were also included as grouped variables by ICD-10 chapter.

2.3. Modelling and Validation

R software which is a free software environment for statistical computing, and R's Classification and Regression Training (caret)[13] and associated packages have been used in our modelling.

Different learning algorithms were implemented, including Random Forests (RF), Linear Discriminant Analysis (LDA), Logistic Regression (LR), Support Vector Machine (SVM), K−Nearest Neighbor (KNN), Elastic Net (ENET), and Neural Network (NN).

The data set was split into a training and a test data set. The training data set consisted of 75% of the cases and a 10-fold cross validation was implemented for the training of the models.

Various standard statistical measures were used for validating the model performance, including sensitivity, specificity, accuracy, Cohen's Kappa, and AUROC.

3. Results

3.1. Data characteristics

We have analyzed the relationship of each factor with delirium to understand their influence on delirium prediction. The following figures show a selection of highly influential factors.

Figure 1 illustrates that the age and comorbidity are two major risk factors for delirium. The median age of the delirium cohort was significantly higher than the median age of the non-delirium cohort, i.e. our control group. That means the higher the age of the patient the more susceptible to develop delirium. Likewise, Figure 1 b) shows that the higher the Charlson comorbidity index was, the higher the probability of developing delirium.

Figure 2 shows that the prevalence of delirium was higher in severely ill patients suffering from other diseases, such as dementia, depression, heart failure, pneumonia or respiratory insufficiency. Nearly 50% of the delirium patients in our cohort group had heart failure, i.e. ~35% of heart failure patients in our considered population have

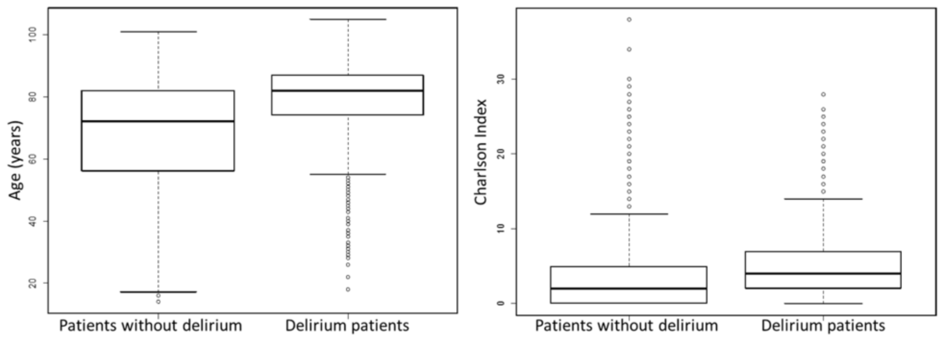

Figure 1 Box plots of (a) age distribution and (b) Charlson comorbidity index distribution for patients in the delirium and non-delirium cohorts

developed delirium. Along with illness, we have also found other factors such as metabolic imbalance, physical disorders etc., through laboratory results and nursing assessment. Examples are presented in Figure 3.

The classification models used along with performance criteria are listed in Table 1. All classification models have shown similar behavior with variation except K-Nearest Neighbors algorithm which was outperformed by the other models.

Figures 4 show the graphical representation of the results for Random Forests.

4. Discussion

The present paper describes how we developed and validated a multivariable prediction model for the occurrence of delirium in hospitalized gerontopsychiatry and internal medicine patients.

Delirium is a very common condition in hospitalized patients. Our paper presents an approach to early-detect patients at risk for delirium based solely on data that already are available at KAGes hospitals. To our knowledge, it is the first publication on predicting delirium with such a large patient population and feature set and our results

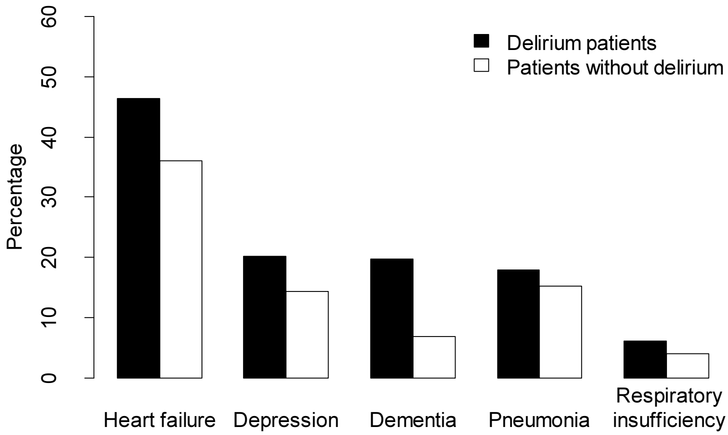

Figure 2 Bar chart showing the percentage of patients with delirium and without delirium w.r.t individual diagnosed disease

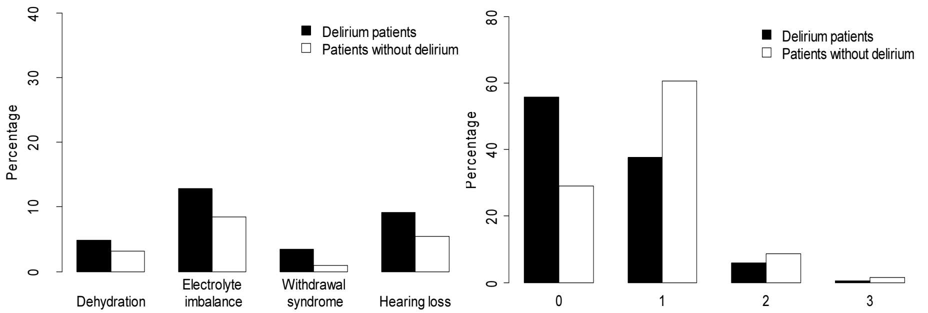

Figure 3 Bar chart showing the significance of metabolic imbalance and physical disorders related to delirium (left). Example of laboratory results, C - reactive protein (CRP) count which is a categorical factor where 0 is normal and 3 is very abnormal (i.e highly elevated) (right).

show excellent prediction accuracy as compared to previous publications. Therefore, we expect that implementation of our models in routine care has the potential to avoid delirium and prolonged hospital stays which would be to the benefit of patients while simultaneously reducing healthcare costs.

4.1. Limitations

Although the model performs very well, there may still be some room for improvement, in particular as far as the collection of the data is concerned.

- From our experience, we expect, that the time and date of delirium occurrences might not always be correct and that in some cases delirium might neither be diagnosed nor recorded.
- The quality of the data used in this work was limited, since the data were not taken from scientific studies featuring dedicated processes for data quality improvement (e.g. source data verification, monitoring, etc.), but the data were taken directly from the KAGes HIS. Therefore, we expect that model accuracy could further be improved, if additional efforts would be applied to data management.
- Although our dataset is large in comparison with previous research, further data with true positive results (i.e. more patients who delivered delirium) would be necessary. We randomly chose 7,000 patients for our control group, while there were only 2,221 delirium patients included. Therefore, a more balanced dataset might further improve our models.
- For implementing our algorithms in clinical routine, higher sensitivity would be advantageous.

Table 1. Comparison of model performance for different modelling approaches

Model	Accuracy	Kappa	Sensitivity	Specificity
Random forest	0.84	0.60	0.69	0.90
Linear discriminant analysis	0.84	0.60	0.69	0.90
Logistic regression	0.82	0.56	0.67	0.88
Support vector machine	0.84	0.61	0.70	0.90
K-nearest neighbor	0.74	0.17	0.13	1.00
Elastic net	0.84	0.59	0.68	0.90
Neural network	0.83	0.60	0.75	0.86

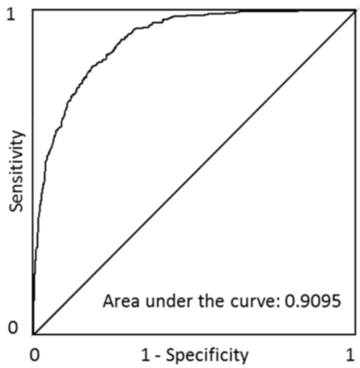

Figure 4. Receiver Operating Characteristics of the Random Forest model with the corresponding performance criteria on the right hand side

4.2. Future work

Our models are still in a preliminary stage so far. Further work on data pre-processing, data-cleaning, model optimization etc. will be necessary before our models can be applied in routine care.

Feature importance needs to be identified to reduce the size of the feature set. Our target is to increase the sensitivity factor without compromising accuracy.

Our analyses revealed significantly different results when we applied different learning algorithms. Further analyses are required to understand the significant model parameters that influence model accuracy.

Acknowledgements

This work has been carried out with the K1 COMET Competence Center CBmed, which is funded by the Federal Ministry of Transport, Innovation and Technology (BMVIT); the Federal Ministry of Science, Research and Economy (BMWFW); Land Steiermark (Department 12, Business and Innovation); the Styrian Business Promotion Agency (SFG); and the Vienna Business Agency. The COMET program is executed by the FFG. KAGes and SAP provided significant resources, manpower and data as basis for research and innovation.

References

[1] T. G. Fong, S. R. Tulebaev, and S. K. Inouye, "Delirium in elderly adults: diagnosis, prevention and treatment," Nat Rev Neurol, vol. 5, no. 4, pp. 210-20, Apr, 2009.

[2] D. K. Kiely, E. R. Marcantonio, S. K. Inouye, M. L. Shaffer, M. A. Bergmann, F. M. Yang, M. A. Fearing, and R. N. Jones, "Persistent delirium predicts greater mortality," J Am Geriatr Soc, vol. 57, no. 1, pp. 55-61, Jan, 2009.

[3] D. M. Popeo, "Delirium in older adults," Mt Sinai J Med, vol. 78, no. 4, pp. 571-82, 2011 Jul-Aug, 2011.

[4] S. Wass, P. J. Webster, and B. R. Nair, "Delirium in the elderly: a review," Oman Med J, vol. 23, no. 3, pp. 150-7, Jul, 2008.

[5] H. Zhang, Y. Lu, M. Liu, Z. Zou, L. Wang, F. Y. Xu, and X. Y. Shi, "Strategies for prevention of postoperative delirium: a systematic review and meta-analysis of randomized trials," Crit Care, vol. 17, no. 2, pp. R47, Mar, 2013.

[6] Y. C. Chin, G. C. Koh, Y. K. Tay, C. H. Tan, and R. A. Merchant, “Underdiagnosis of delirium on admission and prediction of patients who will develop delirium during their inpatient stay: a pilot study,” Singapore Med J, vol. 57, no. 1, pp. 18-21, Jan, 2016.

[7] S. T. O'Keeffe, and J. N. Lavan, “Predicting delirium in elderly patients: development and validation of a risk-stratification model,” Age Ageing, vol. 25, no. 4, pp. 317-21, Jul, 1996.

[8] A. Wassenaar, M. van den Boogaard, T. van Achterberg, A. J. Slooter, M. A. Kuiper, M. E. Hoogendoorn, K. S. Simons, E. Maseda, N. Pinto, C. Jones, A. Luetz, A. Schandl, W. Verbrugghe, L. M. Aitken, F. M. van Haren, A. R. Donders, L. Schoonhoven, and P. Pickkers, “Multinational development and validation of an early prediction model for delirium in ICU patients,” Intensive Care Med, vol. 41, no. 6, pp. 1048-56, Jun, 2015.

[9] J. A. Martinez, A. Belastegui, I. Basabe, X. Goicoechea, C. Aguirre, N. Lizeaga, I. Urreta, and J. I. Emparanza, “Derivation and validation of a clinical prediction rule for delirium in patients admitted to a medical ward: an observational study,” BMJ Open, vol. 2, no. 5, 2012.

[10] M. W. Newman, L. C. O'Dwyer, and L. Rosenthal, “Predicting delirium: a review of risk-stratification models,” Gen Hosp Psychiatry, vol. 37, no. 5, pp. 408-13, 2015 Sep-Oct, 2015.

[11] S. K. Inouye, Y. Zhang, R. N. Jones, D. K. Kiely, F. Yang, and E. R. Marcantonio, “Risk factors for delirium at discharge: development and validation of a predictive model,” Arch Intern Med, vol. 167, no. 13, pp. 1406-13, Jul, 2007.

[12] M. P. Carrasco, L. Villarroel, M. Andrade, J. Calderón, and M. González, “Development and validation of a delirium predictive score in older people,” Age Ageing, vol. 43, no. 3, pp. 346-51, May, 2014.

[13] Max Kuhn. Contributions from Jed Wing, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang, Can Candan and Tyler Hunt. (2016) caret: Classification and Regression Training. R Package Version 6. 0-73. https://CRAN.R-project.org/package=caret