

RESEARCH ARTICLE

Identification of risk factors for the onset of delirium associated with COVID-19 by mining nursing records

Yusuke Miyazawa¹, Narimasa Katsuta^{2*}, Tamaki Nara^{1,3,4}, Shuko Nojiri^{1,3,4}, Toshio Naito⁵, Makoto Hiki^{6,7}, Masako Ichikawa^{6,8}, Yoshihide Takeshita², Tadafumi Kato⁹, Manabu Okumura⁹, Morikuni Tobita^{1,3,4}

1 Department of Healthcare Innovation, Juntendo University Graduate School of Medicine, Tokyo, Japan, **2** Department of Psychiatry, Juntendo University Faculty of Medicine, Tokyo, Japan, **3** Medical Technology Innovation Center, Juntendo University, Tokyo, Japan, **4** Clinical Research and Trial Center, Juntendo University, Tokyo, Japan, **5** Department of General Medicine, Juntendo University Faculty of Medicine, Tokyo, Japan, **6** Department of Emergency and Disaster Medicine, Juntendo University Faculty of Medicine, Tokyo, Japan, **7** Department of Cardiovascular Biology and Medicine, Juntendo University Faculty of Medicine, Tokyo, Japan, **8** Department of Respiratory Medicine, Juntendo University Graduate School of Medicine, Tokyo, Japan, **9** Tokyo Institute of Technology, Tokyo, Japan

* nkatsuta@juntendo.ac.jp



OPEN ACCESS

Citation: Miyazawa Y, Katsuta N, Nara T, Nojiri S, Naito T, Hiki M, et al. (2024) Identification of risk factors for the onset of delirium associated with COVID-19 by mining nursing records. PLoS ONE 19(1): e0296760. <https://doi.org/10.1371/journal.pone.0296760>

Editor: Mihajlo Jakovljevic, Hosei University: Hosei Daigaku, JAPAN

Received: June 20, 2023

Accepted: December 18, 2023

Published: January 19, 2024

Copyright: © 2024 Miyazawa et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data cannot be shared publicly because the participants did not obtain consent for their data to be shared publicly. Data are available from the independent data administrator of this study group (contact via email: y-kurosawa@juntendo.ac.jp) for researchers who meet the criteria for access to confidential data.

Funding: The author(s) received no specific funding for this work.

Abstract

COVID-19 has a range of complications, from no symptoms to severe pneumonia. It can also affect multiple organs including the nervous system. COVID-19 affects the brain, leading to neurological symptoms such as delirium. Delirium, a sudden change in consciousness, can increase the risk of death and prolong the hospital stay. However, research on delirium prediction in patients with COVID-19 is insufficient. This study aimed to identify new risk factors that could predict the onset of delirium in patients with COVID-19 using machine learning (ML) applied to nursing records. This retrospective cohort study used natural language processing and ML to develop a model for classifying the nursing records of patients with delirium. We extracted the features of each word from the model and grouped similar words. To evaluate the usefulness of word groups in predicting the occurrence of delirium in patients with COVID-19, we analyzed the temporal changes in the frequency of occurrence of these word groups before and after the onset of delirium. Moreover, the sensitivity, specificity, and odds ratios were calculated. We identified (1) elimination-related behaviors and conditions and (2) abnormal patient behavior and conditions as risk factors for delirium. Group 1 had the highest sensitivity (0.603), whereas group 2 had the highest specificity and odds ratio (0.938 and 6.903, respectively). These results suggest that these parameters may be useful in predicting delirium in these patients. The risk factors for COVID-19-associated delirium identified in this study were more specific but less sensitive than the ICDSC (Intensive Care Delirium Screening Checklist) and CAM-ICU (Confusion Assessment Method for the Intensive Care Unit). However, they are superior to the ICDSC and CAM-ICU because they can predict delirium without medical staff and at no cost.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The novel coronavirus disease (COVID-19) was declared a pandemic in 2019 by the World Health Organization (WHO) [1]. In Japan, more than 30,000,000 cases of COVID-19 have been reported since the first confirmed case on January 16, 2020 [2, 3]. This disease causes interstitial pneumonia and severe acute respiratory infections with high infection rates. The typical clinical symptoms of COVID-19 include fever, cough, diarrhea, and fatigue. However, atypical presentations include extrapulmonary involvement, such as gastrointestinal symptoms, multi-organ failure, and nervous system involvement. Moreover, there has been growing awareness of the neuropsychiatric manifestations of COVID-19 [4]; coronavirus affects the brain, leading to neurological symptoms such as delirium.

Delirium (sometimes referred to as an “acute confusional state”) is a common clinical syndrome characterized by disturbed consciousness, cognitive function, or perception, with an acute onset and fluctuating course [5]. Assessment of delirium, the most frequent clinical manifestation of acute brain dysfunction, is particularly important in patients with COVID-19 [6]. A review summarizing the incidence of delirium in patients with COVID-19 in 2021 showed that 29–40% of older adults developed delirium [7]. A recent case series from Wuhan, China, reported that 36% of patients with COVID-19 admitted to the hospital experienced neurological conditions, including altered mental states and ischemic strokes [4]. A retrospective study in Saudi Arabia found that one-quarter of patients with COVID-19 developed “confusion,” and almost 9% experienced seizures [8, 9]. Furthermore, COVID-19-associated delirium has been correlated with unfavorable outcomes, including prolonged hospitalization and elevated mortality rates [10, 11]. Several investigations have reported that delirium is prevalent in patients with COVID-19 and is associated with poor clinical outcomes in Japan [12, 13]. Consequently, the management of patients to avoid the onset of delirium is critical, with an escalating need to prognosticate its manifestation.

Delirium prevention is an important issue in COVID-19 infection; however, prevention strategies are nonpharmacological and resource- and personnel-intensive [14]. Current methods for identifying hospitalized patients at an increased risk of delirium require nurse-administered questionnaires with moderate accuracy [15]. However, it is not always possible to have sufficient medical personnel to implement such methods. In particular, during the early stages of the COVID-19 pandemic, the physical and human costs of the infection prevention measures were significant. Therefore, a method that does not interfere with daily operations is needed to identify patients at risk of delirium [16].

Recent significant advances have been made in machine learning (ML) technology, which has been effective in predicting the onset of various diseases [17–20]. Previous studies have used ML to predict the onset of mental disorders, such as depression and delirium [21–23]. These studies on delirium include predictions using ML with structured data, such as age, sex, and weight [16, 24–26], as well as predictions using unstructured text data provided by medical professionals [27–30]. However, a systematic review of predictive models for delirium found that none have been adequately and systematically evaluated [31]. Hence, we aimed to identify previously unknown risk factors for delirium that could be useful in predicting its onset. To achieve this, we used natural language processing (NLP) and ML techniques to analyze nursing records because they contain detailed information regarding patients’ daily conditions.

Materials and methods

Study design

The study period was from January 2021 to May 2021. This was a single-center, retrospective, observational study. Patients admitted to Juntendo University Hospital during the study period with COVID-19 registered as their disease condition were included. When this study was conducted, Juntendo Hospital (Tokyo, Japan) was an acute care hospital with 1,051 beds. The analysis of medical records was conducted from August 2021 to March 2023.

Data source

In this study, patients with COVID-19 infection were diagnosed using the “Guidelines for the Clinical Management of COVID-19” [32]. Diagnoses were primarily confirmed using polymerase chain reaction (PCR), point-of-care (POCT), and antigen testing. Patient data were extracted from information provided by the Juntendo University Information Center. Any information that could identify individual participants was discarded or anonymized prior to the analysis. The extracted information included patient ID, date of birth, sex, medical history, free-text data, and occupation of the individual who described the text data. Notably, nursing records were identified as those recorded by nurses. Patients who (1) did not receive a confirmed diagnosis of infection via a PCR test and (2) were hospitalized for less than 48 h were excluded. In total, 273 patients were included in the analysis.

Nursing records were written multiple times daily by different nurses. The records may be written in SOAP format or as free text. The SOAP format comprises subjective, objective, assessment, and planning information. Specifically, these included the patient’s condition, nurses’ observations, nurses’ review and assessment of vital signs, treatment, messages from the patient, and other comments from the nurse. Some nursing records only documented administrative information, such as prescriptions, and did not include the patient’s condition. This analysis aimed to uncover hidden risk factors through text mining, focusing on patients’ subjective information and nurses’ subjective comments. Therefore, nursing records that contained subjective information were included in the analysis. Fig 1 illustrates this process.

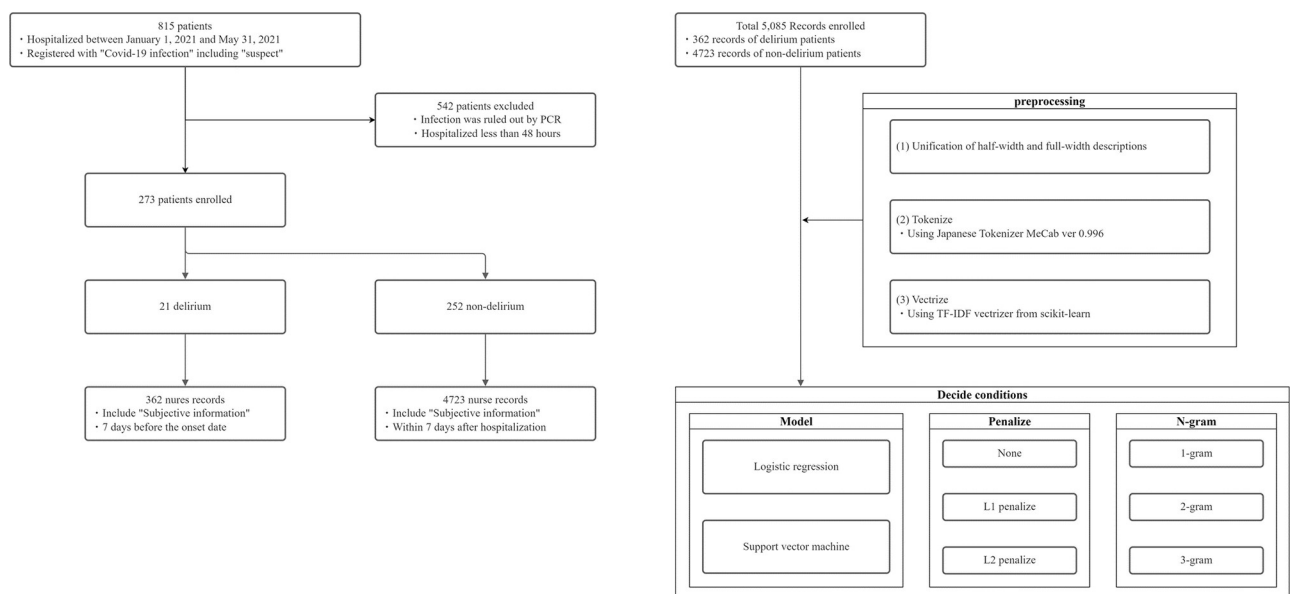


Fig 1. Flow diagram for the records included in this study.

<https://doi.org/10.1371/journal.pone.0296760.g001>

Model development

Natural language texts cannot be used directly by ML algorithms. Therefore, they must be vectorized for processing using ML algorithms. The following are required to perform ML: (1) unification of half- and full-width characters, (2) breaking sentences down to the word level, and (3) vectorization. The text used in this study was written in Japanese. The Japanese language does not have word boundaries. Therefore, it was necessary to break down sentences at the word level. For this task, a Japanese morphological analyzer (MeCab [33]) was used to convert sentences into a standard form of segmented text. This text was then vectorized using the term frequency-inverse document frequency vectorizer from scikit-learn [34]. Through this process, text data written in Japanese were converted into a form that could be processed using ML algorithms. Fig 1 illustrates this process.

ML models were developed to analyze the nursing records of patients with delirium. Logistic regression (LR) [35], support vector machine (SVM) [36], decision tree (DT), and random forest (RF) were used as models. This was implemented using the scikit-learn package in Python [34]. The ML models employed in this study—LR, SVM, DT, and RF—have demonstrated high performance in identifying risk factors for delirium onset in previous studies [37–39]. The primary rationale for selecting these models is their interpretability in influencing the individual factors in the outcome. The objective of the present study was to identify key factors influencing delirium using the trained ML models; therefore, models that enable a clear interpretation of the relationships between factors and the disease are desirable. While some complex ML models, such as neural networks, are known for their high performance, they often pose challenges in interpreting the influence of individual factors. Consequently, the LR, SVM, DT, and RF models were chosen for their capability to quantitatively express the magnitude of the effect of each factor, making these models particularly useful for identifying risk factors for delirium. To optimize the performance of the ML model, we adjusted the parameters and evaluated the results. The model was then evaluated using the four-fold cross-validation method.

Identifying risk factors

The features of each word were extracted from the trained ML model. A total of 50 1-gram words were extracted in the order of feature size. The Japanese language has the following ten parts-of-speech categories: verbs, adjectives, adjectival verbs, nouns, adverbs, participles, conjunctions, impressive verbs, auxiliary verbs, and particles. We defined function words, which include participles, conjunctions, impressive verbs, auxiliary verbs, and particles, as “non-significant words.” Non-significant words and words that appeared fewer than 100 times in the entire text were excluded. Moreover, 1-gram words with similar meanings were categorized into three groups. Subsequently, 2-gram and 3-gram words (in Japanese) were similarly classified into these three groups. This step was based on the opinions of the psychiatrists engaged in the clinical work.

Evaluating risk factors using odds ratios

We identified the risk factors in the preceding steps. Next, we assessed the relationship between these risk factors and the onset of delirium by calculating odds ratios. We aimed to determine whether the words used by each of the three groups could predict delirium. The predictability of the words in each group was evaluated by determining the sensitivity, specificity, and odds ratio of words appearing (1) at least once a day or (2) at least twice a day regarding a particular patient.

Measurement of study variables

The major outcome was delirium incidence. We tracked the presence or absence of delirium and date of onset. Delirium is often not registered in electronic medical records, even if there is an actual onset [40, 41]. Therefore, clinicians at the Juntendo Hospital Mental Clinic identified the onset date of delirium in patients. Experienced psychiatrists diagnosed delirium according to the DSM-5 criteria, based on patients' medical history, physical examinations, and clinical test results, as well as physical and mental status examinations.

Statistical analyses

Patient characteristics are described using the mean (SD) for continuous variables and percentages for dichotomous and categorical variables for all participants, patients with delirium, and patients without delirium. All analyses were performed using Python (version 3.10.5) and scikit-learn (version 1.1.3) [34].

Ethics approval

As this was a retrospective study that dealt with existing medical records, the requirement for informed consent was waived. Information about the study design was posted on the Juntendo University Hospital website, and all candidates were guaranteed the opportunity to refuse to participate in the study. This study adhered to the "Ethical Guidelines for Medical Research for Humans" and the Declaration of Helsinki. The research protocol was approved by the Juntendo University School of Medicine Ethics Committee (approval no. H21-0102).

Results

Study population

The analysis included 273 patients, 21 (7.7%) of whom had delirium. The average age of all patients was 59.8 (SD, 19.1) years, and 61.1% were men. The average age of the patients with delirium was 73.5 (SD, 16.6) years, whereas that of those without delirium was 58.6 (SD, 18.9) years. Moreover, the length of hospital stay was 34.1 (SD, 27.7) and 12.6 (SD, 9.3) days for patients with and without delirium, respectively. The patient characteristics are presented in [Table 1](#).

Construction of the machine learning model

The model was developed using the nursing records of patients with delirium before the onset of delirium and those of patients without delirium as supervisory data. Specifically, the records of patients with delirium up to 7 days before admission were analyzed along with those of patients without delirium within 7 days after admission. The resulting nursing records totaled 5,085. Of these, 362 (7.1%) were nursing records of patients with delirium.

Next, ML models and the N-gram type (1–3-gram) were conditionally trained. The average area under the receiver operating characteristic curve (AUC) was obtained using a four-fold cross-validation method. Using LR, SVM, DT, and RF, the AUC was 0.478–0.788, 0.740–0.804, 0.580–0.647, and 0.758–0.771, respectively. In LR and SVM, the AUC tended to be smaller for L1 regularization. Moreover, no significant differences were observed in the AUC between the 1-gram, 2-gram, and 3-gram models. To identify many possible factor candidates, we decided to use a 3-gram model. According to a systematic review comparing predictive models for the development of delirium in older adults [31], the AUC of the predictive models ranged from 0.52–0.94. The results of the present study were within this range. The results are shown in [Fig 2](#).

Table 1. Demographic characteristics of the patients and nurse records.

	Overall	Delirium	No delirium
N	273	21	252
Male sex	167	13	154
Age (decade, n)			
<10	2	0	2
10 s	7	0	7
20 s	17	1	16
30 s	18	0	18
40 s	32	0	32
50 s	44	5	39
60 s	51	1	50
70 s	59	5	54
80 s	39	6	33
90 s	4	3	1
Past medical history (%)			
Dementia	5	3 (14%)	2 (1%)
Epilepsy	14	1 (5%)	13 (5%)
Depression	10	1 (5%)	8 (3%)
Stroke	21	5 (24%)	16 (6%)
Insomnia	32	4 (20%)	28 (11%)
Hospital day, mean (SD)		34.1 (27.7)	12.6 (9.3)
Nurse record that includes subjective data to be analyzed, n	5085	362	4723

<https://doi.org/10.1371/journal.pone.0296760.t001>

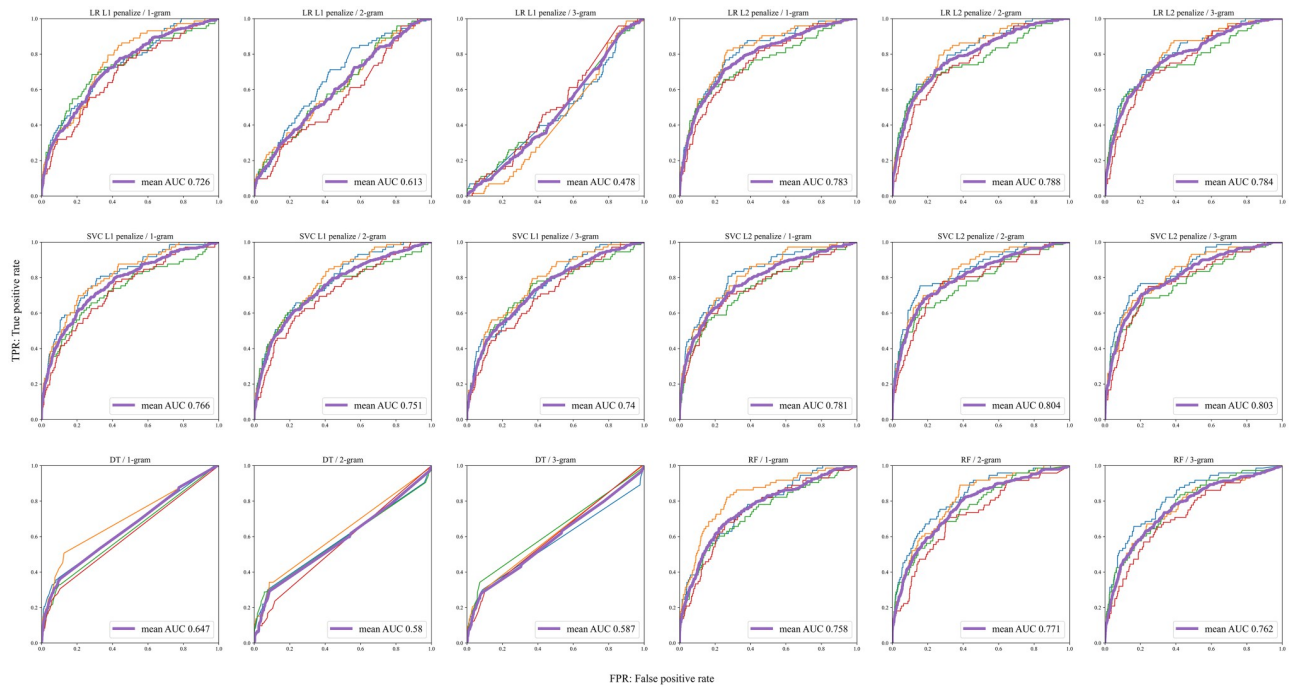


Fig 2. Graphs showing the construction of the machine learning model. LR, logistic regression model; SVC, support vector classifier; DT, decision tree; RF, random forest.

<https://doi.org/10.1371/journal.pone.0296760.g002>

Sensor, Looking After, Talking, Urinating, Needs, Toilet, Nurse Call, Constipation, Calming, Visiting, Understanding, Difficult, Behavior, Saying, Leaving, Decreasing, Danger, Sleeping, Bed, Condition, Assistance, Handling, Attention, Pain, Time, Before/After, Self, Possible, Taking, Wheelchair, Pressure Ulcer, Starting, Choking, Using, Cooling, Body, Insertion, Trouble, Pneumonia, Low, Change, High, Mask, Eating, Meals, Problems, Medications, Moving, Seeing, Behaviors

Fig 3. Fifty 1-gram words with high features were extracted from the machine-learning model.

<https://doi.org/10.1371/journal.pone.0296760.g003>

Identifying risk factors

As described in the Methods section, we identified candidates for risk factors, and factors with similar meanings were consolidated into three distinct groups based on the opinions of the psychiatrists engaged in clinical work: (1) elimination-related behaviors and conditions, (2) abnormal patient behavior and conditions, and (3) subjective expression of pain or severity. A list of words is presented in Fig 3, and the word groups are shown in Fig 4.

Prediction of disease onset using the risk factors

We identified the risk factors in the preceding steps. Next, we evaluated whether these risk factors could predict delirium onset. For each patient, the number of times words in each group appeared per day was tabulated. For each group, we graphed (1) the frequency with which the word appeared at least once and (2) the frequency with which the word appeared at least twice. The sensitivity, specificity, and odds ratio for each group are shown in Fig 5.

Group 1 had the highest sensitivity (0.603). Group 2 had the highest specificity and odds ratio (0.938 and 6.903, respectively). However, group 3 had a low odds ratio (1.985). These results suggest that groups 1 and 2 may be useful in predicting delirium, but group 3 may not.

Group 1. Elimination-related behaviors and conditions

1-gram: Toilet, Urinating, Constipation

2–3-gram: Urine volume, Lead to the restroom, Walk to the restroom, Want to go to the restroom, Urge to urinate, Feeling that bladder has not been completely empty, Dysuria

Group 2. Abnormal patients' behaviors and conditions

1 gram: Sensor, Nurse call, Leaving, Sleeping

2–3-gram: Use/start bed release sensor, Understanding nurse call ambiguously, Describe nurse call

Group 3. Subjective expression of pain or severity

1-gram: Pain, Choking

2–3-gram: Talk about pain, A little choking, Choking oxygen administration, No choking

Fig 4. Grouping of risk factors (excluding non-significant words).

<https://doi.org/10.1371/journal.pone.0296760.g004>

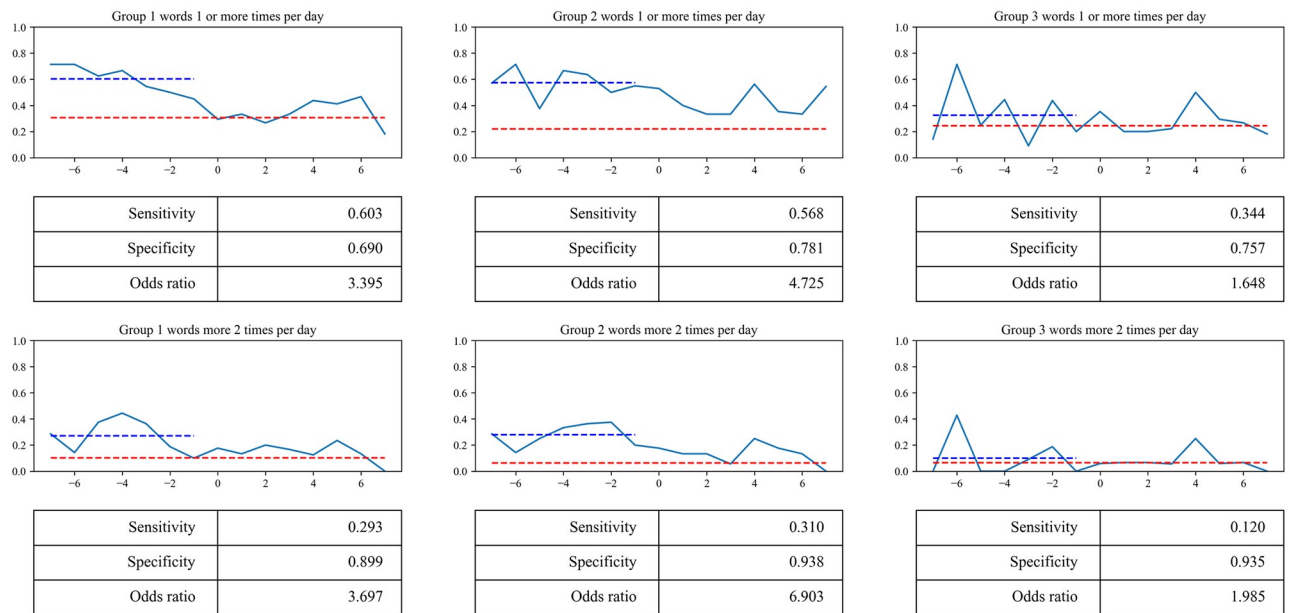


Fig 5. Graphs showing the sensitivity, specificity, and odds ratio for predicting disease onset using risk factors. The vertical axis represents the percentage of patients, and the horizontal axis represents 7 days before and after the day of delirium onset (day 0). The blue dotted line indicates the percentage of patients with delirium 1–7 days before its onset. The red dotted line indicates the frequency observed in patients without delirium.

<https://doi.org/10.1371/journal.pone.0296760.g005>

Discussion

Prior findings

Our objective was to identify unrecognized factors useful for predicting the onset of delirium using ML in the nursing records of patients with COVID-19. With this analysis, we identified (1) elimination-related behaviors and conditions and (2) abnormal patient behavior and conditions as predictors of delirium. In contrast, (3) subjective expression of pain or severity was not a predictor.

However, the identified factors were less sensitive than assessment methods often used in daily practice. Representative assessment methods for delirium include the Intensive Care Delirium Screening Checklist (ICDSC), Confusion Assessment Method for the Intensive Care Unit (CAM-ICU), Delirium Rating Scale-Revised-98 (DRS-R-98), and Memorial Delirium Assessment Scale (MDAS) [42, 43]. In particular, CAM-ICU and ICDSC are often used [44]; CAM-ICU has a sensitivity of 0.771–0.826 and specificity of 0.948–0.968, and ICDSC has a sensitivity of 0.653–0.815 and specificity of 0.767–0.864 [45]. Group 1 demonstrated the highest sensitivity (0.603) but low specificity (0.690). In contrast, group 2 demonstrated the highest specificity (0.938) but low sensitivity (0.310). Compared to using previous methods, prediction using the risk factors identified in this study may not be immediately effective.

However, these assessment methods require additional work for medical staff, and our predictors can be objectively and automatically collected without additional work. Under pandemic conditions, the ability to predict the onset of delirium without additional work and direct contact is desirable, as it can reduce the use of physical resources for infection prevention and risk [16]. In clinical practice, a more detailed evaluation, such as ICDSC, should be performed when these risk factors are identified in patients. This would effectively enable

assessment of the risk of delirium, and necessary management to prevent onset can be performed in advance.

Predictors identified in this study

Group 1, comprising elimination-related behaviors and conditions, included words such as “toilet,” “urination,” and “constipation.” Some studies suggest that elimination is associated with delirium. A lack of physiological control may exacerbate confusion or be considered a symptom. Among the risk assessments for delirium, the NEECHAM Confusion Scale includes an item on “urinary function” [46]. Furthermore, aspects of elimination, such as diaper use, urinary incontinence, and urinary retention, are useful for delirium risk assessment [42, 47, 48]. According to one of the authors, a psychiatric clinician, patients with mental instability occasionally complain of indefinite urinary symptoms. More detailed data collection should be performed in the next step, such as when and how many patients visit the toilet, as it may enable the early prediction of delirium onset.

Group 2, comprising abnormal patients’ behavior and conditions, included words such as “sensor” and “nurse call.” The use of bed sensors or bed fences is a risk factor for delirium development [47, 49], restricting a patient’s body movements [49]. Immobility is thought to be a risk factor, so it encourages medical staff to perform active range-of-motion exercises [5, 42, 48, 50]. To the best of our knowledge, no prior research has been conducted on the relationship between nurse calls and delirium. However, one study statistically summarized the frequency and content of nurse calls, with calls related to the “toilet” being the most common [51]. Given that elimination-related complaints were considered potential risk factors in this study, an increase in nurse calls may reflect these concerns.

Group 3, comprising subjective expression of pain or severity, included words such as “pain” and “choking.” Patients with severe COVID-19 have been reported to develop delirium [4]. Moreover, patient conditions, such as “shortness of breath,” “fever,” “insomnia,” and “catatonic state,” are related to delirium [42, 48, 52–55]. However, in the present study, the relationship between the symptoms and the onset of delirium was weak. One reason for the difference from previous studies may be that the expressions “pain” and “choking” alone do not capture the severity of the symptoms. Objective indices are appropriate for evaluating disease severity. For example, “shortness of breath” should be evaluated using SpO₂ [5, 15]. However, whether scaling subjective symptoms is useful in predicting disease onset remains to be examined.

Change in the frequency of occurrence of predictors before and after delirium onset

We assessed the relationship between these risk factors and the onset of delirium by calculating odds ratios. Fig 5 compares the frequency of occurrence of these factors before and after delirium onset. According to this graph, the frequency of occurrence of words in Groups 1 and 2 was higher before the onset of delirium than in patients without delirium. This demonstrates that the risk factors in Groups 1 and 2 were related to the onset of delirium. However, after the onset of delirium, the frequency decreased and approached that of patients without delirium. We hypothesized that one of the reasons for the decrease after onset may be that therapeutic interventions improved the severity of delirium. Another possible reason for the decrease after onset could be that the therapeutic intervention led to changes in patients’ complaints. For example, if a urinary catheter is inserted as a therapeutic intervention, the frequency of requests to go to the toilet is expected to decrease. Since our study could not examine this aspect, we cannot draw a definitive conclusion.

Strengths and limitations

Our study had several strengths. First, we employed a text mining technique to analyze free-text nursing records. This approach allowed us to successfully identify hidden risk factors such as (1) changes in elimination-related behaviors and conditions and (2) increased frequency of nurse calls and sensor alerts, which were not represented in a structured format in electronic health records. Therefore, the strength of this study lies in its ability to discover findings that cannot be detected by analyzing structured data.

Furthermore, our results can ease the workload of medical staff. Traditional methods, such as ICDSC and CAM, require additional work for patient assessment. For instance, CAM, which is a representative assessment method, requires approximately 5 min [56]. This may overwhelm the staff's daily work if the number of patients is high. In contrast, the risk factors identified in this study can be collected without additional work. For instance, by developing sensors that detect nurse calls or when a patient leaves bed, data can be collected without human intervention. If the collected data meet predefined criteria, alerts can be triggered, thereby identifying patients at high risk of developing delirium before its onset. Early identification facilitates timely intervention, potentially preventing the worsening of outcomes due to delirium. In terms of clinical applications, we envision a product concept of a compact sensor that can fulfill all the necessary functions and be easily attached to the bedside. Such devices can be readily introduced into clinical settings. Moreover, such a device would not increase the chances of contact with patients. During pandemics, factors such as a shortage of human resources make executing standard delirium management practices in routine clinical settings challenging [16, 57]. Another major challenge in the early stages of the COVID-19 pandemic was the risk of spreading infection through contact. We believe that this study offers a solution to this challenge.

However, our study had some limitations. The data used in this study were obtained from a single hospital over a short period. Therefore, our results may not be generalizable. In addition, we used nursing records written in Japanese. Therefore, similar results cannot be obtained from nursing records written in other languages. External validation is required to ensure the generalizability of our results.

Future direction

In this study, we investigated the incidence of delirium. However, NLP and ML have been effective for diseases other than delirium [22, 58]. Additionally, text mining for psychiatric disorders is also effective in assessing depression [21, 58]. Our method may be used to identify predictors of the onset of other psychiatric diseases.

LR, SVM, DT, and RF were used as ML methods. This was done to emphasize the ease of interpreting the results. In general, neural networks and deep learning are effective in improving ML performance [59, 60]. Moreover, Bidirectional Encoder Representations from Transformers have demonstrated high performance in NLP tasks [18, 61]. Ensemble learning, which combines multiple ML methods, is also effective [62, 63]. Furthermore, by adding age, sex, severity of illness, and progress in daily laboratory findings in future research, better prediction performance can be expected.

The data used in this study were limited. An important step in the future would involve applying the predictive keywords and NLP algorithms identified in this study to data from many other hospitals over a long study period. Obtaining a larger dataset from multiple hospitals and sites over a longer study period may help overcome overfitting and bias.

Conclusions

This study successfully employed NLP and ML techniques to analyze nursing records, leading to the identification of specific risk factors for the onset of delirium in hospitalized patients. Our analysis revealed two key predictors: (1) changes in elimination-related behaviors and conditions and (2) increased frequency of nurse calls and sensor alerts. These findings are significant as they can provide a novel, automated approach to delirium risk assessment without imposing additional workload on medical staff. The identified risk factors not only enhance our understanding of early delirium indicators but also offer a practical tool for early intervention, potentially improving patient outcomes in clinical settings.

Acknowledgments

I am grateful to Editage, a service provided by Cactus Communications, for their professional English language editing.

Author Contributions

Conceptualization: Yusuke Miyazawa, Narimasa Katsuta, Tamaki Nara, Shuko Nojiri, Tadamu Kato, Manabu Okumura, Morikuni Tobita.

Data curation: Yusuke Miyazawa, Narimasa Katsuta, Yoshihide Takeshita.

Formal analysis: Yusuke Miyazawa, Manabu Okumura.

Investigation: Yusuke Miyazawa, Narimasa Katsuta, Tamaki Nara, Shuko Nojiri.

Methodology: Yusuke Miyazawa, Narimasa Katsuta, Tamaki Nara, Shuko Nojiri, Manabu Okumura, Morikuni Tobita.

Project administration: Yusuke Miyazawa, Narimasa Katsuta, Tamaki Nara, Shuko Nojiri, Morikuni Tobita.

Resources: Narimasa Katsuta, Toshio Naito, Makoto Hiki, Masako Ichikawa, Yoshihide Takeshita.

Software: Yusuke Miyazawa, Manabu Okumura.

Supervision: Narimasa Katsuta, Tamaki Nara, Shuko Nojiri, Morikuni Tobita.

Validation: Narimasa Katsuta, Tamaki Nara, Shuko Nojiri, Tadamu Kato, Manabu Okumura, Morikuni Tobita.

Visualization: Yusuke Miyazawa, Narimasa Katsuta, Tamaki Nara, Shuko Nojiri, Manabu Okumura.

Writing – original draft: Yusuke Miyazawa.

Writing – review & editing: Yusuke Miyazawa, Narimasa Katsuta, Tamaki Nara, Shuko Nojiri, Toshio Naito, Tadamu Kato, Manabu Okumura, Morikuni Tobita.

References

1. WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020 [cited 2023 May 26]. In: World Health Organization [Internet]. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020>.
2. New Coronavirus Infection Control Number of cases nationwide [cited 2023 May 26]. In: Cabinet Secretariat's Novel Coronavirus Infectious Diseases Control Office. 2023 [Internet]. <https://corona.go.jp/dashboard/>.

3. Current status of the novel coronavirus infection and the response of the MHLW. 4 February 2020 [cited 2023 May 26]. In: Minister of Health, Labour and Welfare [Internet]. https://www.mhlw.go.jp/stf/newpage_09290.html.
4. Mao L, Jin H, Wang M, Hu Y, Chen S, He Q, et al. Neurologic Manifestations of Hospitalized Patients With Coronavirus Disease 2019 in Wuhan, China. *JAMA Neurol.* 2020; 77(6):683–90. <https://doi.org/10.1001/jamaneurol.2020.1127> PMID: 32275288
5. Delirium: prevention, diagnosis and management in hospital and long-term care. 2023 [cited 2023 May 26]. In: NICE | The National Institute for Health and Care Excellence [Internet]. <https://www.nice.org.uk/guidance/cg103>.
6. Arlington, VA: American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed2013.
7. Hawkins M, Sockalingam S, Bonato S, Rajaratnam T, Ravindran M, Gosse P, et al. A rapid review of the pathoetiology, presentation, and management of delirium in adults with COVID-19. *J Psychosom Res.* 2021; 141:110350. <https://doi.org/10.1016/j.jpsychores.2020.110350> PMID: 33401078
8. Saad M, Omrani AS, Baig K, Bahloul A, Elzein F, Matin MA, et al. Clinical aspects and outcomes of 70 patients with Middle East respiratory syndrome coronavirus infection: a single-center experience in Saudi Arabia. *Int J Infect Dis.* 2014; 29:301–6. <https://doi.org/10.1016/j.ijid.2014.09.003> PMID: 25303830
9. Baller EB, Hogan CS, Fusunyan MA, Ivkovic A, Luccarelli JW, Madva E, et al. Neurocovid: Pharmacological Recommendations for Delirium Associated With COVID-19. *Psychosomatics.* 2020; 61(6):585–96. <https://doi.org/10.1016/j.psych.2020.05.013> PMID: 32828569
10. Helms J, Kremer S, Merdji H, Schenck M, Severac F, Clere-Jehl R, et al. Delirium and encephalopathy in severe COVID-19: a cohort analysis of ICU patients. *Crit Care.* 2020; 24(1):491. <https://doi.org/10.1186/s13054-020-03200-1> PMID: 32771053
11. Kennedy M, Helfand BKI, Gou RY, Gartaganis SL, Webb M, Moccia JM, et al. Delirium in Older Patients With COVID-19 Presenting to the Emergency Department. *JAMA Netw Open.* 2020; 3(11):e2029540. <https://doi.org/10.1001/jamanetworkopen.2020.29540> PMID: 33211114
12. Kurahara Y, Matsuda Y, Tsuyuguchi K, Tokoro A. Delirium in Patients with COVID-19 in Japan. *Intern Med.* 2022; 61(19):2861–6. <https://doi.org/10.2169/internalmedicine.9013-21> PMID: 35228428
13. Kandori K, Okada Y, Ishii W, Narumiya H, Maebayashi Y, Iizuka R. Association between visitation restriction during the COVID-19 pandemic and delirium incidence among emergency admission patients: a single-center retrospective observational cohort study in Japan. *Journal of Intensive Care.* 2020; 8(1):90. <https://doi.org/10.1186/s40560-020-00511-x>
14. Martinez F, Tobar C, Hill N. Preventing delirium: should non-pharmacological, multicomponent interventions be used? A systematic review and meta-analysis of the literature. *Age and Ageing.* 2014; 44(2):196–204. <https://doi.org/10.1093/ageing/afu173> PMID: 25424450
15. Wong A, Young AT, Liang AS, Gonzales R, Douglas VC, Hadley D. Development and Validation of an Electronic Health Record-Based Machine Learning Model to Estimate Delirium Risk in Newly Hospitalized Patients Without Known Cognitive Impairment. *JAMA Netw Open.* 2018; 1(4):e181018. <https://doi.org/10.1001/jamanetworkopen.2018.1018> PMID: 30646095
16. Castro VM, Sacks CA, Perlis RH, McCoy TH. Development and External Validation of a Delirium Prediction Model for Hospitalized Patients With Coronavirus Disease 2019. *J Acad Consult Liaison Psychiatry.* 2021; 62(3):298–308. <https://doi.org/10.1016/j.jaclp.2020.12.005> PMID: 33688635
17. Dong J, Feng T, Thapa-Chhetry B, Cho BG, Shum T, Inwald DP, et al. Machine learning model for early prediction of acute kidney injury (AKI) in pediatric critical care. *Crit Care.* 2021; 25(1):288. <https://doi.org/10.1186/s13054-021-03724-0> PMID: 34376222
18. Weng SF, Reys J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One.* 2017; 12(4):e0174944. <https://doi.org/10.1371/journal.pone.0174944> PMID: 28376093
19. Nakatani H, Nakao M, Uchiyama H, Toyoshiba H, Ochiai C. Predicting Inpatient Falls Using Natural Language Processing of Nursing Records Obtained From Japanese Electronic Medical Records: Case-Control Study. *JMIR Med Inform.* 2020; 8(4):e16970. <https://doi.org/10.2196/16970> PMID: 32319959
20. Katsuki M, Narita N, Matsumori Y, Ishida N, Watanabe O, Cai S, et al. Preliminary development of a deep learning-based automated primary headache diagnosis model using Japanese natural language processing of medical questionnaire. *Surg Neurol Int.* 2020; 11:475. https://doi.org/10.25259/SNI_827_2020 PMID: 33500813
21. Vaci N, Liu Q, Kormilitzin A, De Crescenzo F, Kurtulmus A, Harvey J, et al. Natural language processing for structuring clinical text data on depression using UK-CRIS. *Evid Based Ment Health.* 2020; 23(1):21–6. <https://doi.org/10.1136/ebmental-2019-300134> PMID: 32046989

22. Irving J, Patel R, Oliver D, Colling C, Pritchard M, Broadbent M, et al. Using Natural Language Processing on Electronic Health Records to Enhance Detection and Prediction of Psychosis Risk. *Schizophr Bull.* 2021; 47(2):405–14. <https://doi.org/10.1093/schbul/sbaa126> PMID: 33025017
23. Corcoran CM, Carrillo F, Fernández-Slezak D, Bedi G, Klim C, Javitt DC, et al. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry.* 2018; 17(1):67–75. <https://doi.org/10.1002/wps.20491> PMID: 29352548
24. Leslie DL, Inouye SK. The importance of delirium: economic and societal costs. *J Am Geriatr Soc.* 2011; 59 Suppl 2(Suppl 2):S241–3. <https://doi.org/10.1111/j.1532-5415.2011.03671.x> PMID: 22091567
25. Bowman K, Jones L, Masoli J, Mujica-Mota R, Strain D, Butchart J, et al. Predicting incident delirium diagnoses using data from primary-care electronic health records. *Age Ageing.* 2020; 49(3):374–81. <https://doi.org/10.1093/ageing/afaa006> PMID: 32239180
26. Castro VM, Hart KL, Sacks CA, Murphy SN, Perlis RH, McCoy TH Jr., Longitudinal validation of an electronic health record delirium prediction model applied at admission in COVID-19 patients. *Gen Hosp Psychiatry.* 2022; 74:9–17. <https://doi.org/10.1016/j.genhosppsych.2021.10.005> PMID: 34798580
27. Ge W, Alabsi H, Jain A, Ye E, Sun H, Fernandes M, et al. Identifying Patients With Delirium Based on Unstructured Clinical Notes: Observational Study. *JMIR Form Res.* 2022; 6(6):e33834. <https://doi.org/10.2196/33834> PMID: 35749214
28. Gundlapalli AV, South BR, Phansalkar S, Kinney AY, Shen S, Delisle S, et al. Application of Natural Language Processing to VA Electronic Health Records to Identify Phenotypic Characteristics for Clinical and Research Purposes. *Summit Transl Bioinform.* 2008; 2008:36–40. PMID: 21347124
29. Jin H, Wu S. Text Messaging as a Screening Tool for Depression and Related Conditions in Under-served, Predominantly Minority Safety Net Primary Care Patients: Validity Study. *J Med Internet Res.* 2020; 22(3):e17282. <https://doi.org/10.2196/17282> PMID: 32213473
30. Topaz M, Adams V, Wilson P, Woo K, Ryvicker M. Free-Text Documentation of Dementia Symptoms in Home Healthcare: A Natural Language Processing Study. *Gerontology and Geriatric Medicine.* 2020; 6:2333721420959861. <https://doi.org/10.1177/2333721420959861> PMID: 33029550
31. Lindroth H, Bratzke L, Purvis S, Brown R, Coburn M, Mrkobrada M, et al. Systematic review of prediction models for delirium in the older adult inpatient. *BMJ Open.* 2018; 8(4):e019223. <https://doi.org/10.1136/bmjopen-2017-019223> PMID: 29705752
32. Minister of Health, Labour and Welfare. Guidelines for the Clinical Management of COVID-19. <https://www.mhlw.go.jp/content/000936655.pdf>.
33. Kudo T, Yamamoto K, Matsumoto Y, editors. Applying conditional random fields to Japanese morphological analysis. Proceedings of the 2004 conference on empirical methods in natural language processing; 2004.
34. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research.* 2011; 12:2825–30.
35. Stoltzfus JC. Logistic regression: a brief primer. *Acad Emerg Med.* 2011; 18(10):1099–104. <https://doi.org/10.1111/j.1553-2712.2011.01185.x> PMID: 21996075
36. Vapnik V. Estimation of dependences based on empirical data: Springer Science & Business Media; 2006.
37. Lee S, Mueller B, Street WN, Carnahan RM. Machine learning algorithm to predict delirium from emergency department data. *medRxiv.* 2021:2021.02.19.21251956. <https://doi.org/10.1101/2021.02.19.21251956>
38. Mufti HN, Hirsch GM, Abidi SR, Abidi SSR. Exploiting Machine Learning Algorithms and Methods for the Prediction of Agitated Delirium After Cardiac Surgery: Models Development and Validation Study. *JMIR Med Inform.* 2019; 7(4):e14993. <https://doi.org/10.2196/14993> PMID: 31558433
39. Zhao H, You J, Peng Y, Feng Y. Machine Learning Algorithm Using Electronic Chart-Derived Data to Predict Delirium After Elderly Hip Fracture Surgeries: A Retrospective Case-Control Study. *Front Surg.* 2021; 8:634629. <https://doi.org/10.3389/fsurg.2021.634629> PMID: 34327210
40. Han JH, Zimmerman EE, Cutler N, Schnelle J, Morandi A, Dittus RS, et al. Delirium in older emergency department patients: recognition, risk factors, and psychomotor subtypes. *Acad Emerg Med.* 2009; 16(3):193–200. <https://doi.org/10.1111/j.1553-2712.2008.00339.x> PMID: 19154565
41. McCoy TH Jr., Snapper L, Stern TA, Perlis RH. Underreporting of Delirium in Statewide Claims Data: Implications for Clinical Care and Predictive Modeling. *Psychosomatics.* 2016; 57(5):480–8. <https://doi.org/10.1016/j.psych.2016.06.001> PMID: 27480944
42. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet.* 2014; 383(9920):911–22. [https://doi.org/10.1016/S0140-6736\(13\)60688-1](https://doi.org/10.1016/S0140-6736(13)60688-1) PMID: 23992774

43. Hshieh TT, Inouye SK, Oh ES. Delirium in the Elderly. *Psychiatr Clin North Am.* 2018; 41(1):1–17. <https://doi.org/10.1016/j.psc.2017.10.001> PMID: 29412839
44. Patel RP, Gambrell M, Speroff T, Scott TA, Pun BT, Okahashi J, et al. Delirium and sedation in the intensive care unit: survey of behaviors and attitudes of 1384 healthcare professionals. *Crit Care Med.* 2009; 37(3):825–32. <https://doi.org/10.1097/CCM.0b013e31819b8608> PMID: 19237884
45. Gusmao-Flores D, Salluh JI, Chalhub R, Quarantini LC. The confusion assessment method for the intensive care unit (CAM-ICU) and intensive care delirium screening checklist (ICDSC) for the diagnosis of delirium: a systematic review and meta-analysis of clinical studies. *Crit Care.* 2012; 16(4):R115. <https://doi.org/10.1186/cc11407> PMID: 22759376
46. Neelon VJ, Champagne MT, Carlson JR, Funk SG. The NEECHAM Confusion Scale: Construction, Validation, And Clinical Testing. *Nursing Research.* 1996; 45(6):324–30. <https://doi.org/10.1097/00006199-199611000-00002> PMID: 8941300
47. Benedict L, Hazelett S, Fleming E, Ludwick R, Anthony M, Fosnight S, et al. Prevention, detection and intervention with delirium in an acute care hospital: a feasibility study. *Int J Older People Nurs.* 2009; 4(3):194–202. <https://doi.org/10.1111/j.1748-3743.2008.00151.x> PMID: 20925776
48. Solà-Miravete E, López C, Martínez-Segura E, Adell-Lleixà M, Juvé-Udina ME, Lleixà-Fortuño M. Nursing assessment as an effective tool for the identification of delirium risk in older in-patients: A case-control study. *J Clin Nurs.* 2018; 27(1–2):345–54. <https://doi.org/10.1111/jocn.13921> PMID: 28631875
49. Kanno M, Doi M, Kubota K, Kanoya Y. Risk factors for postoperative delirium and subsyndromal delirium in older patients in the surgical ward: A prospective observational study. *PLoS One.* 2021; 16(8): e0255607. <https://doi.org/10.1371/journal.pone.0255607> PMID: 34339463
50. Martins S, Fernandes L. Delirium in elderly people: a review. *Front Neurol.* 2012; 3:101. <https://doi.org/10.3389/fneur.2012.00101> PMID: 22723791
51. Majima Y, Tamura Y, Maekawa Y, Ohira M, Ikegawa M. Survey of actual conditions of reasons for nurse calls: comparing a surgical ward and a medical ward. *Stud Health Technol Inform.* 2009; 146:419–23. PMID: 19592878
52. Ahmed S, Leurent B, Sampson EL. Risk factors for incident delirium among older people in acute hospital medical units: a systematic review and meta-analysis. *Age Ageing.* 2014; 43(3):326–33. <https://doi.org/10.1093/ageing/afu022> PMID: 24610863
53. Enomoto R, Lee-Hiraiwa E. Elimination of the Causes of Poor Sleep Underlying Delirium is a Basic Strategy to Prevent Delirium. *Curr Mol Pharmacol.* 2021; 14(2):132–7. <https://doi.org/10.2174/1874467213666200424150709> PMID: 32329703
54. Chen L, Li N, Zheng Y, Gao L, Ge N, Xie D, et al. A novel semiautomatic Chinese keywords instrument screening delirium based on electronic medical records. *BMC Geriatrics.* 2022; 22(1):779. <https://doi.org/10.1186/s12877-022-03474-w> PMID: 36192690
55. Ceppi MG, Rauch MS, Sándor PS, Gantenbein AR, Krishnakumar S, Albert M, et al. Detecting Incident Delirium within Routinely Collected Inpatient Rehabilitation Data: Validation of a Chart-Based Method. *Neurol Int.* 2021; 13(4):701–11. <https://doi.org/10.3390/neurolint13040067> PMID: 34940753
56. Waszynski CM. How to try this: detecting delirium. *Am J Nurs.* 2007; 107(12):50–9; quiz 9–60. <https://doi.org/10.1097/01.NAJ.0000301029.87489.35> PMID: 18049062
57. Kotfis K, Williams Roberson S, Wilson J, Pun B, Ely EW, Jeżowska I, et al. COVID-19: What do we need to know about ICU delirium during the SARS-CoV-2 pandemic? *Anaesthesiol Intensive Ther.* 2020; 52(2):132–8. <https://doi.org/10.5114/ait.2020.95164> PMID: 32419438
58. Zhou L, Baughman AW, Lei VJ, Lai KH, Navathe AS, Chang F, et al. Identifying Patients with Depression Using Free-text Clinical Documents. *Stud Health Technol Inform.* 2015; 216:629–33. PMID: 26262127
59. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw.* 2015; 61:85–117. <https://doi.org/10.1016/j.neunet.2014.09.003> PMID: 25462637
60. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015; 521(7553):436–44. <https://doi.org/10.1038/nature14539> PMID: 26017442
61. Strobl AN, Vickers AJ, Van Calster B, Steyerberg E, Leach RJ, Thompson IM, et al. Improving patient prostate cancer risk assessment: Moving from static, globally-applied to dynamic, practice-specific risk calculators. *J Biomed Inform.* 2015; 56:87–93. <https://doi.org/10.1016/j.jbi.2015.05.001> PMID: 25989018
62. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *Lancet Respir Med.* 2015; 3(1):42–52. [https://doi.org/10.1016/S2213-2600\(14\)70239-5](https://doi.org/10.1016/S2213-2600(14)70239-5) PMID: 25466337
63. Rose S. Mortality Risk Score Prediction in an Elderly Population Using Machine Learning. *American Journal of Epidemiology.* 2013; 177(5):443–52. <https://doi.org/10.1093/aje/kws241> PMID: 23364879