Supporting the early detection of disease onset and change using document vector analysis of nursing observation records

Shotaro Komaki[1,2], Fuminori Muranaga[3],

Yumiko Uto[4], Takashi Iwaanakuchi[1,3], Ichiro Kumamoto[4]

[1]Kagoshima University Graduate School of Medical and Dental Science, Kagoshima, Japan

[2]Kagoshima Medical Professional College, Kagoshima, Japan

[3]Medical Informatics, Kagoshima University Hospital, Kagoshima, Japan

[4]Department of Medical Information Sciences, Graduate School of Medicine and Dental Sciences, Kagoshima University, Kagoshima, Japan

Corresponding author:

Shotaro Komaki (MMSc)

Kagoshima University Graduate School of Medical and Dental Science

8-35-1, Sakuragaoka, Kagoshima Kagoshima 890-8544 Japan

E-mail: st.komaki@harada-gakuen.ac.jp

Phone number: +81-99-275-5176

Fax number: +81-99-275-5177

Supporting the early detection of disease onset and change using document vector analysis of nursing observation records

Abstract

Nursing records are an account of patient condition and treatment during their hospital stay. In this study, we developed a system that can automatically analyze nursing records to predict the occurrence of diseases and incidents (e.g., falls). Text vectorization was performed for nursing records and compared with past case data on aspiration pneumonia, to develop an onset prediction system. Nursing records for a patient group that developed aspiration pneumonia during hospitalization and a non-onset control group were randomly assigned to definitive diagnostic (for learning), preliminary survey, and test datasets. Data from the preliminary survey were used to adjust parameters and influencing factors. The final verification used the test data and revealed the highest compatibility to predict the onset of aspiration pneumonia (sensitivity = 90.9%, specificity = 60.3%) with the parameter values of size = 80 (number of dimensions of the sentence vector), window = 13 (number of words before and after the learned word), and min_count = 2 (threshold of wordcount for word to be included). This method represents the foundation for a discovery/warning system using machine-based automated monitoring to predict the onset of diseases and prevent adverse incidents such as falls.

Key words: Narrative Medicine; Patient Safety; Aspiration Pneumonia; Machine Learning; Natural Language Processing

Introduction

Nursing observation records

In health care institutions, one must always be aware of the condition of one's patients. Among medical professionals, nurses have the closest daily contact with patients, and are most likely to notice and record changes in the conditions of patient. Nurses carefully monitor patients, maintain records on their current conditions, and predict possible future changes in their conditions (Imai and Takase, 2017; Sasaki et al., 2019). Their records are a valuable source of information to other medical staff. Unlike records on biochemical tests, medical imaging tests, and other indicators, nursing observation records are an accumulation of information that provides an up-to-date view of the condition of the patient from the professional perspective of nurses and healthcare professionals. These records are an important source of information for the rapid identification of changes in the condition of a patient. In addition, although nursing record writing has been published widely in many journals, there are only slight variations in nursing record content among authors (Muramatsu et al., 2010). Because the nursing record contents contain text that has a structural element, they could be adapted for machine learning as well. Nurses sometimes face the challenge of deciding whether to report the observed changes in the status of a patient to doctors or their superiors (Kunjukunju and Ahmad, 2019). Thus, changes often noted in nursing records are not reported to doctors or superiors. Considering these cases, we propose a system that can extract high-risk patients based on an objective assessment of nursing records and contribute to the decision-making process for intervention by medical professionals.

Aspiration pneumonia

Pneumonia is the third most common cause of death in Japan (Ministry of Health and Welfare, n.d.; Statistics Bureau Ministry of Internal Affairs and Communications, n.d.). The most common type of pneumonia among elderly patients is aspiration pneumonia, which increases the mortality rate during hospitalization, lengthens hospital stays, and introduces additional medical costs (Hitoshi, 2016; Ingeman et al., 2011; Wilson, 2012). Preventing the onset of aspiration pneumonia is important, but because the status of patients changes throughout their hospitalization, the current approach of screening for risk of aspiration pneumonia at the time of admission only is insufficient. Furthermore, 20–30% of elderly patients, who represent the main patient population affected by this condition, do not exhibit typical symptoms, such as cough, sputum, fever, and dyspnea, making the diagnosis difficult (Son et al., 2017). Thus, a more efficient method of early onset detection is required to reduce the number of patient deaths, length of hospital stays, and medical costs.

Introducing natural language processing in disease onset prediction

Recently, various methods for natural language processing focused on unstructured records have been studied. We previously analyzed nursing observation records using term frequency–inverse document frequency (TF–IDF) to predict aspiration pneumonia (Shotaro et al., 2015). We successfully extracted the characteristic vocabulary for the onset of disease; however, we were unable to markedly differentiate between aspiration pneumonia and other diseases. Vector encoding is used to place components of natural language, such as the words in a sentence, in an N-dimensional space and represent them as vectors so that they can be handled computationally. In the present study, we used the paragraph vector method, advocated by Le and Mikolov (2014), to investigate text vectorization methods for textual data. One approach to implement the Paragraph Vector method is the use of the Doc2Vec tool of the Gensim library (Řehůřek and Sojka, 2010). Doc2Vec is an application version of the Word2Vec tool developed by Mikolov et al. (2013). Word2Vec uses the one-hot vector of a word as the input layer, weights it in the intermediate layer, and outputs the updated value as a word vector when propagating the error from the output layer to the intermediate layer (Hirokazu, 2014). Doc2Vec uses the same premise but for sentence analysis, and can be used to determine the similarity by comparing sentence vectors.

Doc2Vec has a large number of parameters that must be set in advance. To test the concept, we used the default parameters of Doc2Vec (i.e., parameters used in the analysis example of the Gensim library) to test a small amount of patients records to determine if aspiration pneumonia could be detected. Even when using the default parameters, it was possible to remarkably differentiate aspiration pneumonia and control cases.

Our trials revealed that the identification of the most appropriate parameter ranges is necessary for the optimal analysis of nursing observation records (Shotaro et al., 2017, 2018). Furthermore, some limitations were observed in previous studies associated with the use of non-unified text amounts and a small number of datasets. Therefore, in the present study, we unified the amount of text used, increased the number of cases, and determined the most appropriate range of parameters. Furthermore, we applied the determined parameters to other datasets, verified the accuracy of the proposed method, and evaluated whether it can be used as an effective system to predict the onset of diseases.

Purpose

The purpose of this study is to develop a system that can analyze nursing observation records and predict the onset of aspiration pneumonia. Using a Doc2Vec-based text vectorization technique, the optimal range of parameters suitable for analyzing nursing observation records for prediction was identified and evaluated.

Method

Analysis records

Subjects

Aspiration pneumonia cases were extracted from records of patients who were admitted to the Kagoshima University Hospital between 2013 and 2017, diagnosed with aspiration pneumonia, and administered with antibiotics. Were comprehensively reviewed the medical records, drug histories, and laboratory results of the extracted cases with the help of the collaborating physicians, to obtain clinically definitive diagnoses of aspiration pneumonia. Records of patients admitted between June 1 and July 20, 2018 were extracted as control cases. The inclusion criteria for the control group consisted of records of patient of age corresponding to the 95% confidence interval of the aspiration pneumonia case group.

Structure of the analyzed records and pre-processing

The nursing observation records analyzed were limited to 40 characters (80 bytes) per record, so that the condition of each patient was sequentially recorded using simple expressions without redundancy (Chiaki et al., 2009). In this analysis, the number of records over a period of four days was set as the verification target; records obtained until the onset date were used for the aspiration pneumonia case group, and records from the hospitalization date were used for the control case group.

Sentences in Japanese do not contain spaces between words; thus, it was necessary to insert spaces between words by performing a morphological analysis as pre-processing. In this study, the morphological analysis was performed using MeCab (Kudo et al., 2004). In addition to the Information-Technology Promotion Agency Dictionary (IPADIC) attached to MeCab, ComeJisyo (ComeJisyo, n.d.), in which single-letter words are deleted, and the user dictionary developed by the Kagoshima University Hospital, were employed as dictionaries for the morphological analysis. Nursing observation records unrelated to either the aspiration pneumonia case group or the control group data were used to create the user dictionary; words that are not properly divided (e.g., proper nouns) were extracted and registered as part of the user dictionary.

Analysis and evaluation methods

Aspiration pneumonia cases were randomly assigned to three groups to create three datasets: definitive diagnostic dataset (for learning), preliminary survey dataset, and test dataset. The definitive diagnostic dataset was derived only from aspiration pneumonia cases; the data from control cases were added to the preliminary survey and test datasets. Cosine similarity was calculated for the definitive diagnostic dataset by applying the preliminary survey and test datasets to the Doc2Vec model in individual succession. In other words, the similarity of the definitive diagnostic dataset, consisting of previous aspiration pneumonia case records, to each of the other datasets was determined based on cosine similarity. Because the cosine similarity exhibited a variation even within the same dataset, repeated trials were implemented to calculate the average cosine similarity of the output data for learning, and these were set as the cosine similarities for each of the datasets.

This study was conducted by setting three parameters from the preliminary survey data: number of definitive diagnostic data records (for learning) used for calculations, number of iterations, and search parameter ranges. We then used the test dataset to verify the accuracy of the final parameters.

Preliminary survey

The preliminary survey was conducted to establish the above-mentioned three parameters. Each parameter is explained in this section.

All definitive diagnostic data were output using cosine similarity. In this study, we used the mean average precision (MAP) as an index to determine the appropriate number of data records (TopN) in the survey.

When performing an analysis using a neural network and a tool such as Doc2Vec, a weighting adjustment is made with random numbers to improve the calculation efficiency (Kouki, 2016). As a result, the cosine similarity changes slightly with each iteration, and repeated trials were conducted to find the average value. Specifically, the cosine similarity of the document was calculated using the definitive diagnostic data and preliminary survey data for the aspiration pneumonia cases. To verify the number of iterations before and after, the number of trials at which the variance converged was verified using an F-test.

In this study, the size (10–100), window (4–15), and min_count were adjusted as the search parameters. The average precision (AP) of each parameter was used to determine the optimal search range. The size can be defined as the number of dimensions of the sentence vector, the window as the number (context) of words that are used in a sentence before and after the word used for learning, and min_count as a threshold employed to discard words with lower appearance frequencies (Gensim topic modelling for humans, n.d.). Five aspiration pneumonia cases were randomly sampled from the definitive diagnostic and preliminary survey datasets, and the number of appearances of the feature vocabulary extracted from TF–IDF was used to determine the upper limit of min_count (1).

$$\text{TF-IDF}(i)=\text{TF}(i)\times\text{IDF}(i)=\text{TF}(i)\times\text{Log2}(N/\text{DF}(i)) \quad (1)$$

Ethical concerns

This study was reviewed and approved by the head of the Ethics Committee on Epidemiological Research, Graduate School of Medical and Dental Sciences, Kagoshima University (Number: 697).

Results

Preliminary survey data

In total, 98 aspiration pneumonia cases and 228 control cases within the target time periods were extracted and used as data for analysis. The analysis showed that certain terms occurred only once in 47–83% of the records. The physical size of the individual records was 2–16 kB for the aspiration pneumonia group, and 1–10 kB for the control group. By erroneous deglutition 98 cases of pneumonia were bifurcated: 49 cases comprised the confirmed diagnosis data (for learning), and the other 49 cases comprised the preliminary survey data. Therefore, all the definitive diagnosis data comprised aspiration pneumonia cases. Additionally, 228 control subjects were included in the preliminary study data.

The appropriate number of records used for the definitive diagnostic dataset (for learning) calculation (TopN) was verified. TopN denotes the upper limit of the number of records used for calculation. The analysis showed that the top nine records, with a MAP of 0.45, deliver the most accurate results. This was followed in descending order by the top four (MAP = 0.43), top three (MAP = 0.42), and top five (MAP = 0.40) records. When all 49 records were used, the lowest MAP value of 0.22 was noted.

Show the Number of dispersion convergences. The dispersion converged between 8 and 9 iterations (p = 0.24).

The TF–IDF values for five randomly sampled aspiration pneumonia cases were calculated. The top 5 words for each case are listed in Table 1. The vocabulary in the table is presented in the original Japanese form, followed by phonetic representation in English script and the corresponding translations (letters and abbreviations are written verbatim).

[Insert Table 1 near here]

Estimating the optimal parameter range using the preliminary survey dataset

The survey ranges were determined to verify the optimal parameter values. The min_count of the definitive diagnostic data and the preliminary survey data indicated that 47–83% of the terms appear only once in each record. In addition, the TF–IDF results demonstrated that even words from the vocabulary list in Table 1 that appear only two or three times in a sentence are important (e.g., in Case 4: Loperamide = 2 times, pyrexia = 3 times). Therefore, the upper limit of min_count was set to 3. The size was set to 10–100, and the window was set to 4–15, based on a previous study (Shotaro et al., 2017, 2018). Additionally, min_count was set to 1–3, and the search parameter range was determined in this study.

Based on the optimal parameter verification results, the number of iterations was set to 8, and the top 9 records were set for the definitive diagnostic dataset used for calculations. As a result of the search parameter range identification described above, the top 10 highest AP values from the calculated parameters are listed in Table 2.

[Insert Table 2 near here]

The parameter values that maximized the AP values were of size 40, window of 12, and min_count of 1. The area under the curve (AUC) in the corresponding receiver operating characteristic (ROC) curve was 0.718, and the maximum sensitivity and specificity were determined to be 81.6% and 57.9%, respectively.

Verification of the test dataset

Size and basic attributes of the test dataset

In total, 33 aspiration pneumonia cases and 63 control cases within the target time periods were used in the test dataset. The analysis showed that even in the test dataset, certain terms occurred only once in 51–83% of the records. The individual nursing observation records of the aspiration pneumonia group were 2–18 kB in size, while those of the control group were 1–12 kB in size.

Optimal parameter ranges for the test dataset

The number of iterations, training dataset records used for calculations, and search parameter ranges were determined under the same conditions as those of the preliminary survey data. Table 3 presents the top 10 AP scores for each calculated parameter value.

[Insert Table 3 near here]

The optimal parameter values corresponding to the highest AP for the test dataset were a size of 80, window of 13, and min_count of 2. The AUC in the corresponding ROC curve was 0.763, and the maximum sensitivity and specificity were determined to be 90.9% and 60.3%, respectively.

Discussion

Size of definitive diagnostic dataset (for learning) for calculation and parameter search ranges

The top nine cases in the definitive diagnostic dataset (for learning) exhibited the highest MAP; the lowest MAP was found when all cases were used (top 49). As aspiration pneumonia can be caused by various factors, such as food aspiration during meals and silent aspiration during sleep (Kohno et al., 2013), the content of nursing observation records may not be very consistent. According to the analysis results, inclusion of all records of definitive diagnostic data in the calculations resulted in the emergence of various patterns. To improve accuracy, rather than using all data examples to detect aspiration pneumonia, only the top nine cases with high similarity were used.

The search parameter of size refers to the number of vector dimensions. Mikolov et al. (2013) stated that if the number of dimensions of a feature vector is doubled, the number of words used for learning should also be doubled. In our previous study, we found that a size of 10 enables the most remarkable differentiation when using seven cases for the definitive diagnostic dataset (Shotaro et al., 2018). In this study, the dataset was expanded to include 49 cases, with 40–100 dimensions accounting for the majority in the top 10 AP (Tables 2 and 3). Therefore, the number of vector dimensions must be increased in proportion as the dataset increases. The window is the number of words (context) before and after the word used for learning in a sentence. The optimal range refinement was a difficult parameter to set in our previous studies (Shotaro et al., 2017, 2018). In the present study, the window width in the top 10 AP scores mostly consisted of values of 10 or greater (Tables 2 and 3), which can be attributed to the morphemes being finely divided according to the usage expressions in Japanese. For example, the Japanese expression "食べていなかった (tabeteinakatta)" can be divided into five morphemes: "食べ (tabe)," "て (te)," "い (i)," "なかっ (nakat)," and "た (ta)." Owing to the nature of the nursing observation records (Muramatsu et al., 2010), simple expressions were used; however, the morphemes were still finely divided. Therefore, it is more efficient to use a wide window for records written in Japanese. Min_count is a threshold value that discards vocabulary with fewer-than-specified number of occurrences. In each record, 47–83% of the terms appeared only once; therefore, if the min_count value is increased, the amount of text would decrease. Moreover, the TF–IDF results of evaluating feature vocabulary demonstrated that even vocabulary that appears only two or three times in a sentence is important (Table 1). Therefore, setting the min_count higher may result in the discarding of important vocabulary. Consequently, when analyzing nursing observation records, it is necessary to set the min_count threshold low.

For best results, the upper limit of the search range for the size parameter should increase as the size of the definitive diagnostic dataset increases. Further, it is necessary to increase the window width and set min_count to a small value when analyzing nursing observation records written in Japanese.

Benefits of using natural language processing analysis for nursing observation records

In this study, we analyzed text to predict the onset of aspiration pneumonia. Text, a form of unstructured data, is more difficult to normalize and classify than structured data, which can be easily processed and sampled. However, the nursing observation records analyzed in this study were written by professional nurses, who are required to use accurate and appropriate expressions. In addition, many articles and specialized journals have reported guidelines on writing these records (Muramatsu et al., 2010); further, many learning opportunities are available to write these records (e.g., training courses and medical settings). Therefore, the content of the descriptions can be expected not to vary markedly between records, i.e., nursing observation records can be expected to have a more uniform writing style than other medical records. The nursing observation records used in this study were also used in a previous study and found to be useful. In the previous study, natural language processing was applied to record audit screening to identify and retrain nurses who used inappropriate expressions in nursing observation records (Chiaki et al., 2010). Moreover, the usefulness of text analysis in the medical field was demonstrated by Goodwin et al. (2017), who analyzed the interview records of mentally handicapped persons and reported that they were useful for inferring the severity score. Therefore, we believe that the analysis of nursing observation records can contribute to decision-making in the medical field and help in the detection of disease onset.

The Doc2Vec used in this study is based on the same principles of Word2Vec (Le and Mikolov, 2014), which allows for the drastic reduction of computational complexity compared with other available methods (Dias et al., 2014; Yan and Zhu, 2018). Therefore, it can be used without excessively overloading the computer system even when analyzing electronic medical records, which contain vast amounts of data (Huang, 2013). In addition, aspiration pneumonia cases could be significantly differentiated even with a small amount of data, eliminating the need for the large amount of data generally used in machine learning. Nursing observation records are useful for ascertaining changes in patient conditions because they do not necessarily require much definitive diagnostic information, such as structured data (e.g., biochemical tests), and they can be evaluated with a high degree of accuracy. Currently, the use of electronic medical records is rapidly becoming popular in Japan, as well as in other countries (Japanese Association of Healthcare Information Systems Industry, 2018; Pedersen et al., 2012; Roland et al., 2012; van Weel et al., 2012). With the progression of electronic management of patient information, an automated monitoring system that is necessary as a medical safety system for predicting the onset of diseases and preventing incidents such as falls could be developed.

Limitations of the study

Although this study has many advantages, the analysis of nursing observation records alone is not sufficient to determine the signs of the onset of aspiration pneumonia. However, various risk factors have been cited for the development of aspiration pneumonia (van der Maarel-Wierink et al., 2011), and by adding the analysis score of this method to the overall judgment, we expect to improve the accuracy of predicting aspiration pneumonia.

Also, Patients prone to developing aspiration pneumonia are in the chronic phase of the disease (Son et al., 2017), but university hospitals in Japan function such that only a small proportion of wards are set aside for chronic-phase patients (Ministry of Health and Welfare, n.d.). Therefore, this situation made it difficult to collect a large amount of data in this study. Studies with small sample sizes are inherently likely to have inflated effect estimates (Button et al., 2013; Ioannidis, 2018). Therefore, the results of this study may likewise be overestimated and should be carefully examined through the lens of sample size. We believe that future studies can collect sufficient samples jointly from university hospitals and further expand upon this study.

Conclusions

In this study, nursing observation records were effectively analyzed using Doc2Vec to develop an onset prediction system. A system that is capable of automatically monitoring nursing observation records is necessary in the current era of electronic management of patient information. Aspiration pneumonia is known to be the major cause of death in elderly patients, and the occurrence of this disease can result in increased medical costs and the use of antibiotics (Hitoshi, 2016; Ingeman et al., 2011; Wilson, 2012). The purpose of this study was to predict the onset of aspiration pneumonia by refining the investigation scope of effective Japanese nursing observation record parameters. The developed system could be utilized to predict the onset of aspiration pneumonia and other diseases and prevent the occurrence of incidents such as falls. The technique adopted in this study is expected to contribute to the prevention of aspiration pneumonia and benefit both patients and medical caregiving institutions.

Acknowledgment

We would like to thank the personnel at the Department of Medical Informatics, Kagoshima University Hospital, who helped us obtain the data used in this study.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience.* DOI: 10.1038/nrn3475.

Chiaki, T., Fuminori, M., & Yumiko, U. (2010). Nurse's record audit using text mining analysis. *30th Joint Conference on Medical Informatics,* 30, pp. 572–575, Hamamatsu, Japan.

Chiaki, T., Mukaikubo, S., Yumiko, U. & Miyuki, K. (2009*).* Refinement of nurse's record interface for narrative note in electronic medical, *29th Joint Conference on Medical Informatics 29,* pp. 410–413, Hiroshima, Japan.

ComeJisyo (n.d.) Comejisyo project top page - OSDN. https://ja.osdn.net/projects/comedic/ (accessed 25 March 2017).

Dias, A., Gorzelniak, L., Schultz, K., Wittmann, M., Rudnik, J., Jörres, R., & Horsch, A. (2014). Classification of exacerbation episodes in chronic obstructive pulmonary disease patients. *Methods of Information in Medicine*, *53*(2), 108–114. https://doi.org/10.3414/ME12-01-0108

Genism topic modelling for humans (n.d.) models.word2vec – Word2vec embeddings. https://radimrehurek.com/gensim/models/word2vec.html (accessed 10 February 2017).

Goodwin, T. R., Maldonado, R., & Harabagiu, S. M. (2017). Automatic recognition of symptom severity from psychiatric evaluation records. *Journal of Biomedical Informatics*, *75*, S71–S84. https://doi.org/10.1016/j.jbi.2017.05.020

Hirokazu, N. (2014). *Word2vec Ni Yoru Shizen Gengo Syori Nyuumon (Natural Language Processing by Word2vec)*. Tokyo: O'Reilly Japan.

Hitoshi, O. (2016). Additional cost in medical care and length of stay due to aspiration pneumonia that inpatients developed: Observational study using DPC data. *The Journal of the Japan Society for Health Care Management*, *17*(3), 123–128.

Huang, Y.C. (2013). Mining association rules between abnormal health examination results and outpatient medical records. *Health Information Management Journal*, *42*(2), 23–30. https://doi.org/10.1177/183335831304200204

Imai, T., & Takase, M. (2017). Support system in the workplace to improve the quality of nursing practice by newly graduated nurses: Analyzing data collected through interviews with nursing managers, using the KJ method. *Japanese Journal of Occupational Medicine and Traumatology*, *65*(3), 111–117. https://ci.nii.ac.jp/naid/40021212540/en/

Ingeman, A., Andersen, G., Hundborg, H. H., Svendsen, M. L., & Johnsen, S. P. (2011). In-hospital medical complications, length of stay, and mortality among stroke unit patients. *Stroke*, *42*(11), 3214–3218. https://doi.org/10.1161/STROKEAHA.110.610881

Ioannidis, J. P. (2018). Why most published research findings are false. In: *Getting to Good: Research Integrity in the Biomedical Sciences*. DOI: 10.1371/journal.pmed.0020124.

Japanese Association of Healthcare Information Systems Industry (2018). JAHIS Order entry Electronic medical record introduction survey report 2017 edition.

https://www.jahis.jp/action/id=57?contents_type=23 (accessed 27 December 2018).

Kohno, S., Imamura, Y., Shindo, Y., Seki, M., Ishida, T., Teramoto, S., Kadota, J., Tomono, K. & Watanabe, A. (2013). Clinical practice guidelines for nursing- and healthcare-associated pneumonia (NHCAP) [Complete translation]. *Respiratory Investigation*, *51*(2), 103–126. https://doi.org/10.1016/j.resinv.2012.11.001

Kouki, S. (2016). *Zero Kara Tsukuru Deep Learning : Python de Manabu Deep Learning No Riron to Zissen (Make It from Scratch Deep Learning - Learning with Python Theory and Implementation of Deep Learning)*. Tokyo: O'Reilly Japan.

Kudo, T., Yamamoto, K., & Matsumoto, Y. (2004). *Applying conditional random fields to Japanese morphological analysis.* Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, Barcelona, Spain. https://pdfs.semanticscholar.org/62d8/219e6be879d54b5602c90fdbeaff7a78ff74.pdf

Kunjukunju, A., & Ahmad, A. (2019). Effective communication among doctors and nurses: Barriers as perceived by doctors. *The Malaysian Journal of Nursing*, *11*(2), 3–11. https://doi.org/10.31674/mjn.2019.v11i02.001

Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. *Proceedings of the 31st International Conference on Machine Learning*, 32(2), 1188–1196. https://arxiv.org/pdf/1405.4053.pdf

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*, arXiv preprint arXiv:1301.3781. https://arxiv.org/abs/1301.3781

Ministry of Health and Welfare (n.d.) Statistics and Information Department. Vital Statistics 2016. http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/suikei16/dl/2016suikei.pdf (accessed 27 January 2017).

Ministry of Health and Welfare (n.d.) The 9th Study Group on Guidelines for Formulating Regional Medical Initiatives, etc. https://www.mhlw.go.jp/file/05-Shingikai-10801000-Iseikyoku-Soumuka/0000078

123.pdf (accessed March 5, 2021).

Muramatsu, H., Watanabe, I., Osaki, C. & Kazuto, K. (2010). Text mining of nursing records. *Information Processing Society of Japan TOD*, *3*(3), 112–122.

Pedersen, K.M., Andersen, J.S., & Sondergaard, J. (2012). General practice and primary health care in Denmark. *The Journal of the American Board of Family Medicine*, *25*(Suppl 1), S34–S38. https://doi.org/10.3122/jabfm.2012.02.110216

Řehůřek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. *Proceedings of the LREC Workshop on New Challenges for NLP Frameworks*, Valetta, Malta. https://www.fi.muni.cz/usr/sojka/papers/lrec2010-rehurek-sojka.pdf

Roland. M., Guthrie, B., & Thome, D.C. (2012). Primary medical care in the United Kingdom. *The Journal of the American Board of Family Medicine*, *25*(Suppl 1), S6–S11. https://doi.org/10.3122/jabfm.2012.02.110200

Sasaki, M., Uto, Y., Yoshida, T., Iwaanakuchi, T., Muranaga, F., Saigo, Y., & Kumamoto, I. (2019). Secondary use of hospital information system data for safe bedside radiography in terms of patient factors. *Health Information Management Journal*, *48*(1), 24–32. https://doi.org/10.1177/1833358317749162

Shotaro, K., Fuminori, M., Hiroshi, K., Miho, Y., Takashi, I., Yumiko, U. & Yukari F. (2015) *Efforts to predict the development of aspiration pneumonia from matters described in medical records, and the evaluation thereof* - *Utilizing data mining* -. The 34th Joint Conference on Medical Informatics 34: 750–753, Chiba, Japan.

Shotaro, K., Fuminori, M., Takashi, I., Yumiko, U. & Ichiro, K. (2017) *Prevention of aspiration pneumonia by sentence vectorization technique.* The 37th Joint Conference on Medical Informatics 37: 380–383, Osaka, Japan.

Shotaro, K., Fuminori, M., Takashi, I., Yumiko, U. & Ichiro, K. (2018) *Evaluation of aspiration pneumonitis discovery method by sentence vectorization of observed records of nurse.* Japan Association for Medical Informatics Spring Conference: 134–135, Niigata, Japan.

Son, Y.G., Shin, J., & Ryu, H.G. (2017). Pneumonitis and pneumonia after aspiration. *Journal of Dental Anesthesia and Pain Medicine*, *17*(1), 1–12.

https://doi.org/10.17245/jdapm.2017.17.1.1

Statistics Bureau Ministry of Internal Affairs and Communications (n.d.) Statistical Handbook of Japan 2016. https://www.stat.go.jp/english/data/handbook/pdf/2016all.pdf (accessed 25 January 2017).

van der Maarel-Wierink, C.D., Vanobbergen, J.N.O., Bronkhorst, E.M., Schols, J.M., & de Baat, C. (2011). Risk factors for aspiration pneumonia in frail older people: A systematic literature review. *Journal of the American Medical Directors Association, 12*(5), 344–354. https://doi.org/10.1016/j.jamda.2010.12.099

van Weel, C., Schers, H., & Timmermans, A. (2012). Health care in The Netherlands. *The Journal of the American Board of Family Medicine, 25*(Suppl 1), S12–S17. https://doi.org/10.3122/jabfm.2012.02.110212

Wilson, R.D. (2012). Mortality and cost of pneumonia after stroke for different risk groups. *Journal of Stroke and Cerebrovascular Diseases, 21*(1), 61–67. https://doi.org/10.1016/j.jstrokecerebrovasdis.2010.05.002

Yan, E., & Zhu, Y. (2018). Tracking word semantic change in biomedical literature. *International Journal of Medical Informatics, 109*, 76–86. https://doi.org/10.1016/j.ijmedinf.2017.11.006

Table 1. Words with the highest Term Frequency-Inverse Document Frequency (TF-IDF) values for five randomly selected aspiration pneumonia cases. (V = Vocabulary; N = Number)

Case 1		Case 2		Case 3		Case 4		Case 5	
V	N	V	N	V	N	V	N	V	N
L	11	嘔吐 “outo” [Vomiting]	8	吸引 “kyuuin” [Suction]	7	ロペミン “ropemin” [Loperamide]	2	エネーボ “eneebo” [Enevo]	6
分 “hun” [Minute]	14	上田 “ueda” ※A person's name	3	褥瘡 “zyokusou” [Bedsore]	5	熱 “netsu” [Fever]	3	大声 “oogoe” [Loud]	6
酸素 “sanso” [Oxygen]	10	SPO2	8	瞳孔不同 “doukouhudou” [Anisocoria]	4	低値 “teichi” [Low value]	3	入眠 “nyuumin” [Sleeping]	15
5	13	多量 “taryou” [A lot]	6	多量 “taryou” [A lot]	6	配薬 “haiyaku” [Medication]	2	多量 “taryou” [A lot]	10
/	14	L	6	白色痰 “hakusyokutan” [White sputum]	4	下痢 “geri” [Diarrhea]	3	ホリゾン “horizon” [Horizon]	4

Table 2. Top 10 average precision (AP) values for various parameter values (preliminary survey data)

size	window	min_count	AP
40	12	1	0.436205426
60	6	1	0.433256461
40	8	1	0.430546809
80	12	1	0.428175516
60	13	1	0.422276537
60	14	1	0.418351127
40	11	1	0.416334529
40	12	2	0.414355221
100	12	1	0.414306112
70	15	1	0.408291844

Table 3. Top 10 average precision (AP) values for various parameter values (test dataset)

size	window	min_count	AP
80	13	2	0.572954618
90	12	1	0.563011535
100	12	2	0.562416931
10	11	3	0.559600684
70	12	3	0.558893377
80	12	2	0.558722658
30	12	2	0.55675763
80	10	2	0.556043434
80	15	2	0.555136424
90	12	2	0.555121959