Patterns in Stroke Occurrence on Warm Days in Winter by Associations Analysis

Hiroshi Morimoto

Graduate School of Environmental Study, Nagoya University, Nagoya, Japan
Email: hiroshim@info.human.nagoya-u.ac.jp

Abstract

There is increasing interest in the effects of meteorological factors on stroke occurrence. A number of studies have found that cold exposure was a trigger to the onset of cerebral infarction in winter. It was believed that the case of stroke occurrence on warm days in winter was just a coincidence and negligible. The main objective of this study was to offer the evidence against the above belief. We aimed to find typical patterns among stroke occurrences in case the weather became warmer in winter. Associations analysis was conducted using daily data of emergency transport at Nagoya City in Japan. Typical patterns were identified among risky stroke occurrences even for the case that the weather became warmer in winter. These findings will contribute to an accurate forecast of stroke occurrence in winter.

Keywords

Stroke, Cerebral Infarction, Weather, Associations Analysis

1. Introduction

Numerous studies investigated the effects of meteorological factors on the occurrence of many diseases such as stroke or cerebral infarction ([1] Field *et al*. 2002, [2] Ebi *et al*. 2004, [3] Jimenez-Conde *et al*. 2008). It was reported that there was very close negative correlation between air temperature and the mortality rates for cerebral infarction ([4] Bull and Morton 1975). Cold exposure caused increases in mortality from ischaemic heart disease, too ([5] Donaldson and Keating 1997). These evidences suggested that if the temperature became colder, then the risk of cerebral infarction would increase. According to these ideas, a weather watch/warning system for stroke had been carried out in various cities (e.g., at Hamburg in Germany with Berichtuber die Tatigkeit, at Hiroshima prefecture in Japan with Home page of the association of doctors in Hiroshima). They set some threshold temperature and then provided warning or alarm if the temperature fell below that threshold.

In the background of distinction of the above approaches, there was the tacit consent that the cases of the risky days with relatively warm weather in winter were simply a coincidence and negligible. The main objective of this study was to try to show that the increase of patients of stroke on relatively warm days in winter was not a result of randomness but a result of certain rule. We aimed to find a typical rule in these risky warm days, using association rules mining.

In method section, we introduced association analysis. Results section described the processing of the original data for association analysis. The findings of the typical patterns were given by association rule mining. In discussion section, we argued some possible reasons why the cases of warmer days in winter became risky against cerebral infarction.

2. Method

2.1. Data

In this paper, we used three types of data: the data of emergency transport, the weather data and the weather pattern data. The data of emergency transport was the daily data from 2002 January up to 2004 December in Nagoya City in Japan. This data was supplied by Nagoya City Fire Department. The weather data was the daily data of various whether elements, such as temperature (mean, maximum and minimum temperature) and sea-level pressure supplied by Japan Meteorological Agency. The data of weather pattern were extracted from Japan Weather Association (JWA), which were called "11-types of weather patterns" including the pattern of "high pressure", "a cold front" and so on. See for details [6] Kojima *et al.*, 2014, [7] Morimoto, H., 2015.

2.2. Association Rule Mining

Association rules analysis has proved a highly successful technique for extracting useful information from very large databases. We proceed to define the notations of association analysis, mainly based on [8] Ordonez (2006). See also [9] Li (2006) for basics of this method. Let D be a database like a matrix. In our case, D was supposed to be a database of a matrix, where the row labels were dates (from 2002/1/1 to 2004/12/31) and the column labels were the information of the number of patients of cerebral infarction, weather elements and so on. In Ordonez (2006), D was called a set of "transactions". Here "patients" meant those people who were transported by emergency and were diagnosed with cerebral infarction.

Let X be a condition that requires that some value of a column j of D is a given value x for each row:

{the value of the column j of D is equal to the value x}.

X corresponds to "item" in Ordonez (2006). In our case, X is a condition like

{the mean temperature of the day in question is equal to 10}.

Let Y be another condition, like

{the value of the column k of D is equal to y}.

An expression $\{X \Rightarrow Y\}$ is called an association rule, where X is called "antecedent" and Y is called "consequent".

For given X, $P(X)$ is defined as a fraction of the number of rows of D that satisfies the rule X over the total number of D. This is equal to the probability that the rule X is satisfied.

$$P(X) = \#(\text{row of } D \text{ that satisfies } X)/\#(D)$$

where #(set) denotes the number of the set.

Hereafter we identify D as the set of rows and X as the set of rows that satisfies the rule X. By this notation, we can write $X \subset D$.

Association rules are estimated by some useful measures such as "support", "confidence" and "lift". Now let $\{X \Rightarrow Y\}$ be an association rule.

The confidence is defined as the conditional probability of Y given X,

$$conf(X \Rightarrow Y) = P(Y \mid X)$$

This is equal to

$$\#(X \cap Y)/\#(X) = P(X \cap Y)/P(X)$$

The lift is defined by

$$lift\left(X \Rightarrow Y\right) = conf\left(X \Rightarrow Y\right)/P(Y) = \left[P(X \cap Y)\right]/\left[P(X)P(Y)\right]$$

If the lift is equal to one, then the probability of selecting the element of Y is equal to that of Y given X, *i.e.*, the condition X is independent from Y. If the lift is greater than one, then the condition X is meaningful and the rule $\{X \Rightarrow Y\}$ can be estimated to be a good rule. Therefore, "confidence" and "lift" are good measure to quantify association rules.

The support of the rule $\{X \Rightarrow Y\}$ is defined as the probability of $X \cap Y$, and denoted by

$$\sup\left(X \Rightarrow Y\right) = \#(X \cap Y)/\#(D) = P(X \cap Y).$$

Although this definition of support is different from Ordonez (2006), we followed the way of the software R that was used in this paper.

In our case, D was supposed to be a database of a matrix with dates from 2002/1/1 to 2004/12/31 as rows. D had, as columns, risk (number of patients who were diagnosed as cerebral infarction after transport to hospital), temperature, sea-level pressure and so on. The condition Y is something like.

{the level of cerebral infarction is high}.

X might be

{the temperature is very low}

or X might be a combination of two or more conditions like

{The temperature is very low, the sea level pressure is low},

here "," means "and". Then for example, one possible rule could be an association rule like

{The temperature is very low, the sea level pressure is low ≥ the risk of cerebral infarction is high}.

If the confidence and lift are relatively large then the association rule can be regarded as a good rule. Although association rules represent a promising technique, they produce an extremely large number of rules. Most of such rules are irrelevant. To improve this problem, the notions such as "lift" are useful to find practical rules.

3. Results

We processed the original data supplied by Nagoya City Fire Department to fit the association rules mining. The process was performed in two steps.

First, we compared the data of the day in question with the data of the previous day and made a new data with their differences and so on. For example, the data of the day 2002/1/2 was composed of the data with the differences of the day 2002/1/1: the data of 2002/1/2 were defined to be the differences of sea level pressure, mean temperature, maximum temperature, minimum temperature with the date of 2002/1/1, the weather pattern of 2002/1/1, the weather pattern of 2002/1/2 and the number of patients of cerebral infarction who were transported to hospital by emergency. As a result of this process, our data begins with the date 2002/1/2, and the first part of our new data was listed in **Table 1**.

Secondly, we simplified the plus or minus numbers as "up", "down" or "even". Precisely speaking, if the absolute number of sea level pressure were less than or equal to 4, then we replace the numbers by "even". If sea level pressure was greater (or less) than 4, we replaced it by "up" (or "down"). The value "4" was selected because mostly the absolute value was less than 40. Similarly, we changed values of various temperature (such as mean, maxim or minimum) to "up", "down" or "even", if they were greater than 1, less than 1, or the absolute value was less than or equal to 1.

As for weather charts, we noticed that they were basically classified into two patterns, high pressure type and low pressure type. Therefore we abbreviated "highp" and "lowp" for high pressure type and low pressure type.(In fact, we replaced the types 3, 4, 5, 6 to high pressure type and 1, 2, 7, 8, 9, 10, 11 to low pressure type. For these numbers, see Morimoto, H. 2015).For the number of patients, we simplified them as "high risk", "low risk" and "mid risk". The thresholds were determined by calculating distribution of the patients numbers, and were defined so that the ratioof low risk (and high risk too) occupied 25% and mid risk 50%. Thus if the number of patients was less than or equal to 2, then the risk was set the risk "low". If the number of patients was higher than or equal to 7, then we set the risk "high". The others were set "mid". Finally the data were simplified as listed in **Table 2**. The data had 8 columns and 270 rows (daily data in winter during three years from 2002 to

Table 1. Data of weather elements, weather patterns and risks of cerebral infarction.

Date	seapres_d	meantmp_d	maxtmp_d	mintmp_d	Pre-pattern	Pattern	Risk
2002/1/2	−6.6	−1.4	−3.4	−2.7	10	4	6
2002/1/3	6.5	−2.9	−4.6	−0.4	4	4	7
2002/1/4	−0.3	2	2.9	1.3	4	5	7
2002/1/5	−2.8	1	1.9	1.6	5	4	5
2002/1/6	12.2	−0.6	0.5	−1.4	4	5	10

Here, "seapres_d" means the difference of sea level pressure with the previous day and so on. "pre-pattern" and "pattern" mean the weather chart numbers of the previous day and the day in question. "risk" implies the number of patients who were diagnosed as cerebral infarction after transport to hospital.

Table 2. Data arranged for associations analysis.

Date	seapres_d	meantmp_d	maxtmp_d	mintmp_d	Pre-pattern	Pattern	Risk
2002/1/2	down	down	down	down	lowp	highp	mid
2002/1/3	up	down	down	even	highp	highp	high
2002/1/4	even	up	up	up	highp	highp	high
2002/1/5	even	up	up	up	highp	highp	mid
2002/1/6	up	even	even	down	highp	highp	high

Here, "lowp" (or "highp") expressed that the weather chart was of low (or high) pressure type.

2004).

Associations analysis:

In this paper, association rules mining was performed by a software "R" with the package called "arules". Typical patterns were identified among risky stroke occurrences even in case that the weather became warmer in winter. Since the whole rules found were extremely large, we only searched those rules of "support" and "confidence" greater than 0.02 and 0.7. There were totally 8 rules (as listed **Table 3**) with "consequence {risk= high}", *i.e.*, the conditions that deduced high risk of cerebral infarction.

Some common features of high risk for cerebral infarction were obviously observed in the list of **Table 3**. The risky days were identified on the condition that sea level pressure decreased, temperature increased and at the same time, weather chart changed from high pressure state to low pressure state. These findings suggested that there existed characteristic pattern among risky days on warmer days in winter.

We proceeded to investigate further what the association rules in **Table 3** suggested. The research was done to find the concrete days that really satisfied the conditions (*i.e.*, antecedent) in winters from 2002 to 2004 in our database. The typical days were listed in **Table 4**. The details of these results were discussed in the following section.

4. Discussion

Association analysis in the previous section revealed the existence of some typical weather conditions that triggered the onset of cerebral infarction (see **Table 3**). These conditions contained the followings in common:

1) The temperature (mean, maximum or minimum) increased.

2) Sea-level pressure decreased.

3) The weather pattern of the previous day was mostly good and the area was covered with high atmospheric pressure.

4) The weather pattern of the day in question was not really good and the area was covered with low atmospheric pressure.

The conditions (2)-(4) were quite natural as the conditions for the occurrence of cerebral infarction, because the weather had tendency to become worse. The condition (1) might seem to be contradictory against (2)-(4). To resolve this contradiction, we investigated in detail the days that matched to conditions (1)-(4), *i.e.*, these days

Table 3. List of association rules.

Antecedent	Consequent	Support	Confidence	Lift
seapres_d = down, meantmp_d = up, previouspat = highp, pattern = lowp	risk = high	0.022	0.857	3.127
seapres_d = down, meantmp_d = up, previouspat = highp, pattern = lowp	risk = high	0.022	0.857	3.127
meantmp_d = up, previouspat = highp, pattern = lowp	risk = high	0.033	0.818	2.985
meantmp_d = up, mintmp_d = up, previouspat = highp, pattern = lowp	risk = high	0.033	0.818	2.985
seapres_d = down, meantmp_d = up, pattern = lowp	risk = high	0.026	0.778	2.838
seapres_d = down, meantmp_d = up, mintmp_d = up, pattern = low	risk = high	0.026	0.778	2.838
maxtmp_d = even, mintmp_d = up, previouspat = highp, pattern = lowp	risk = high	0.022	0.75	2.736
meantmp_d = up, mintmp_d = up, pattern = lowp	risk = high	0.037	0.714	2.606

Table 4. risky days that appeared in the association rules.

Date	Sea-Level Pressure	Mean Temp.	Max- Temp.	Min- Temp.	Previous Pattern	Present Pattern	Risk	Patients
2002/1/16	down	up	up	up	highp	lowp	high	7
2002/1/21	down	up	up	up	lowp	lowp	high	7
2002/12/19	even	up	even	up	highp	lowp	high	10
2002/2/22	down	up	even	up	highp	lowp	high	10
2003/1/27	down	up	up	up	highp	lowp	high	7
2003/2/10	even	up	even	up	highp	lowp	high	13
2003/2/23	down	up	up	up	highp	lowp	high	7
2003/12/6	down	up	down	up	highp	lowp	high	11
2004/12/2	up	up	even	up	highp	lowp	high	8
2004/12/4	down	up	even	up	highp	lowp	high	7

listed in **Table 4**. For this purpose, the information of weather chart or map of these days was extracted from Japan Meteorological Agency.

For example, we checked the day 2002/2/22 that commonly appeared in the **Table 4** and the previous day 2002/2/21 too. We showed the weather charts of these days in **Figure 1**. These charts will be explored in more detail in the following paragraph.
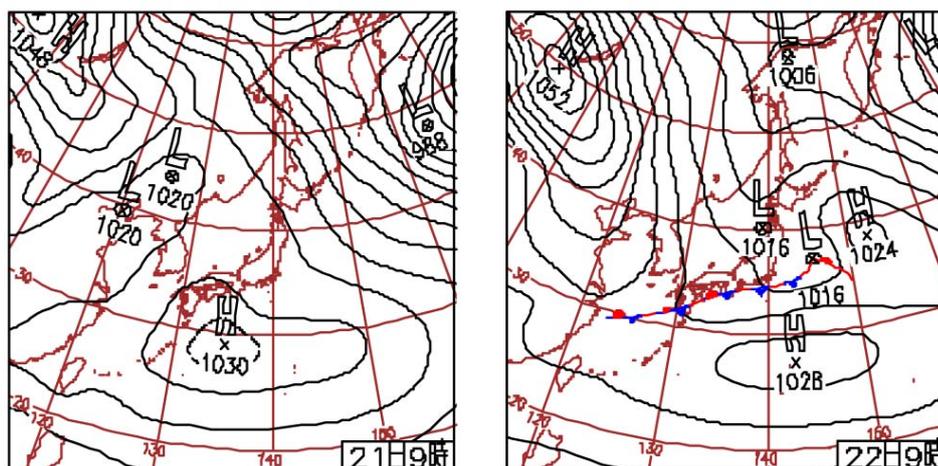
Figure 1. Weather Charts of 2002/2/21(left) and 2002/2/22(right).

On the day 2002/2/21, the whole area of Japan was covered with a air of high atmospheric pressure and the weather was good. As for temperature, Japan Meteorological Agency announced the following comments. "The temperature was high like in April. Small cabbage white butterfly (Pierisrapae) was observed for the first time." But the weather chart of 2002/2/22 showed the existence of a stationary front extending in the south sea. It was rainy in Nagoya. Japan Meteorological Agency commented that "some dandelion blossomed one month earlier than normal years in Osaka". This comment, together with that of 2002/2/21, showed that the temperature was kept relatively high in the background during the two days.

These evidences suggested that there existed a high risk of cerebral infarction when the weather pattern shifted to worse state but the temperature had been kept to be high and became even higher in the day in question.

From a medical point of view, a reasonable explanation for this finding could be provided by an effect of nervous system. It is well known that human's health is maintained by a good balance of the sympathetic nervous system and the parasympathetic nervous system. If there is a stimulus or impact against Human's body (such as exposure to cold weather), then human's nervous system acts first. And then, an internal secretion begins to work ([10] Heindle *et al*., 2004). The increase of temperature and the deterioration of weather were contradictory. This contradiction might have caused complexity or disturbance in the balance of human's nervous system. It is well known disturbance of the autonomic nervous system cause serious disease. The contradictory condition of weather might become a trigger to the onset of cerebral infarction. As for hospital admissions, it was reported that the effects of weather conditions occurred some days after exposure, e.g., exposure to cold air ([11] Pablo *et al*. 2009). But in our cases, the temperature was kept high and the weather was maintained to be relatively warm. Therefore, the increase of cerebral infarction was not the type of an effect of delay.

5. Conclusions

It was believed that cold exposure was a trigger to the onset of cerebral infarction in winter. The case of stroke occurrence on warm days in winter had been neglected as an accidental exception. We aimed to find whether there existed a typical rule in the increase of patients of cerebral infarction on relatively warm days in winter or not. We applied the theory of association rule mining to the daily data of emergency transport at Nagoya City in Japan. This associations analysis revealed the existence of typical patterns among risky stroke occurrences for the case that the weather became even warmer in winter.

As a result, we can conclude that there is a pattern to the onset of cerebral infarction in relatively warm days in winter: If

1) as a background, relatively warm days continue,
2) the sea level pressure decreases,
3) the weather chart changes from high pressure type to low pressure type,
4) the temperature increases,

then, the risk of cerebral infarction increases.

This discovery will contribute to an improvement of watch/warning system of cerebral infarction in winter.

Acknowledgements

References

[1] Field, T.S. and Hill, D. (2002) Weather, Chinook, and Stroke Occurrence. *Stroke*, **33**, 1751-1758. http://dx.doi.org/10.1161/01.STR.0000020384.92499.59

[2] Ebi, K.L., Exuzides, K.A., Lau, E., Kelsh, M. and Barnston, A. (2004) Weather Changes Associated with Hospitalizations for Cardiovascular Diseases and Stroke in California, 1983-1998. *International Journal of Biometeorology*, **49**, 48-58. http://dx.doi.org/10.1007/s00484-004-0207-5

[3] Jimenez-Conde, J., Ois A., Gomis, M., Rodriguez-Campello, A, Cuadrado-Godia, E, Subirana, I. and Roquer, J. (2008) Weather as a Trigger of Stroke. *Cerebrovascular Diseases*, **26**, 348-354. http://dx.doi.org/10.1159/000151637

[4] Bull, G.M. and Morton, J. (1975) Seasonal and Short-Term Relationship of Temperature with Deaths from Myocardial and Cerebral Infarction. *Age and Ageing*, **4**, 19-31. http://dx.doi.org/10.1093/ageing/4.1.19

[5] Donaldson, G.C. and Keatinge, W.R. (1997) Early Increases in Ischaemic Heart Disease Mortality Dissociated from and Later Changes Associated with Respiratory Mortality after Cold Weather in South East England. *Journal of Epidemiology & Community Health*, **51**, 643-648. http://dx.doi.org/10.1136/jech.51.6.643

[6] Kojima, M., Ogasawara, N. and Morimoto, H. (2014) Relationship between Onset of Cerebral Infarction and Classification of Meteorological Elements. *Open Journal of Applied Sciences*, **4**, 242-248. http://dx.doi.org/10.4236/ojapps.2014.45024

[7] Morimoto, H. (2015) Study on Links between Cerebral Infarction and Climate Change Based on Hidden Markov Models. *International Journal of Social Science Studies*, **3**, 180-186. http://dx.doi.org/10.11114/ijsss.v3i5.1045

[8] Ordonez, C. (2006) Association Rule Discovery with the Train and Test Approach for Heart Disease Prediction. *IEEE Transactions on Information Technology in Biomedicine*, **10**, 334-343. http://dx.doi.org/10.1109/TITB.2006.864475

[9] Li, J. (2006) On Optimal Rule Discovery. *IEEE Transactions on Knowledge and Data Engineering*, **18**, 460-471. http://dx.doi.org/10.1109/TKDE.2006.1599385

[10] Heindl, S., Struck, J., Wellhöner, P., Sayk, F. and Dodt, C. (2004) Effect of Facial Cooling and Cold Air Inhalation on Sympathetic Nerve Activity in Men. *Respiratory Physiology & Neurobiology*, **142**, 69-80. http://dx.doi.org/10.1016/j.resp.2004.05.004

[11] Pablo, F., Tomas, C., Soriano, L.R. and Diego, L. (2009) Winter Circulation Weather Types and Hospital Admissions for Cardiovascular, Respiratory and Digestive Diseases in Salamanca, Spain. *International Journal of Climatology*, **29**, 1692-1703. http://dx.doi.org/10.1002/joc.1802