Analysis of news of the Japanese asbestos panic: a supposedly resolved issue that turned out to be a time bomb

Yoshimitsu Takahashi¹,², Koichi Miyaki¹, Takeo Nakayama¹

¹Department of Health Informatics, Kyoto University School of Public Health, Yoshida Konoe, Sakyo, Kyoto, Japan 606-8501
²Department of Clinical Research and Informatics, International Clinical Research Center, Research Institute, International Medical Center of Japan, 1-21-1 Toyama, Shinjuku, Tokyo, Japan 162-8655
Address correspondence to T. Nakayama, E-mail: nakayama@pbh.med.kyoto-u.ac.jp

ABSTRACT

Background Asbestos-linked public health problems were widely reported in Japan, in 2005. The objective is to apply text mining with network analysis to characterize these problems.

Methods Text mining with network analysis of newspaper headlines including the word ‘asbestos’ published in 1987 and 2005 was conducted. Outcome measures are occurrence of the words and simultaneous occurrence of two words in the newspaper headlines.

Results In 36 headlines, which contained the word ‘asbestos’ in 1987, the word ‘pollution’ (40%) appeared most frequently, followed by ‘removal’ (31%) and ‘campaign’ (29%). For combinations of words, the following occurred most frequently: ‘campaign and expulsion’ (26%) followed by ‘removal and campaign’ (14%). Of 293 headlines in 2005, the following words appeared: ‘hazard’ (31%), ‘person’ (16%) and ‘death’ (13%). For combinations, the following appeared: ‘person and death’ (9%). Asbestos pollution and removal campaigns were reported in 1987, but the death of citizens was reported in 2005.

Conclusions Text mining with network analysis, which presents one of the methods for visualization of text data, suggests the following insight. Insufficient steps against asbestos had been taken for 20 years, which is compatible with the latency period. It has resulted in widespread exposure to asbestos and more severe asbestos-related public health problems among citizens. This methodology suggests that analyzing text data by this method can serve future surveillance and efficient use of epidemiological knowledge.

Keywords public health, statistical methods

Introduction

In the summer of 2005, asbestos-linked deaths were reported in Osaka. Some individuals were employed by a farm equipment manufacturer, and two individuals lived near the manufacturing facility. This first report of asbestos-related deaths among the general public caused severe concern in Japan and demands for specific measures. The outcry forced the Ministry of Health, Labor and Welfare (MHLW) to declare a total ban on the use of asbestos by 2008.

The well-established risks associated with asbestos

The risks associated with the use of asbestos are well documented. The risks associated with two specific kinds of asbestos, crocidolite (blue asbestos) and amosite (brown asbestos), respectively, were reported in 1972 by the International Labor Organization (ILO) and the International Agency for Research on Cancer (IARC). The risk of chrysotile (white asbestos) was reported by the IARC in the 1980s, and an asbestos convention was upheld by the ILO in 1986. The IARC then concluded that asbestos was carcinogenic to humans and classified it as a Group 1 carcinogen.¹-⁴ A characteristic of asbestos-linked disease is that there is a long latency period; a latency period of lung cancer is at least 10–20 years.⁵,⁶ By the 1980s, most developed countries had prohibited or placed restrictions on the use of asbestos.⁷ As a result of these prohibitions, the predicted peak years of
mesothelioma related to asbestos use are 2004 for Europe and 2015 through 2020 for the United States. However, the peak in Japan will be somewhat later, 2025, because the total ban on the use of asbestos was not implemented during the 1980s (Fig. 1).8

The Japanese public first became aware of the dangers of asbestos in existing construction in 1987. Reports that asbestos had been used in the construction of schools created widespread awareness of the risks associated with the use of the material. Although the risks related to the use of asbestos had become well known, the importation of asbestos continued to increase in the late 1980s.9 In the past 20 years, the health hazards of asbestos have been well established and accepted, and the long latency period has been documented epidemiologically and medically, but little of the reports about asbestos were written; after 20 years of silence about asbestos, awareness of this issue was raised in the summer of 2005 with the deaths of the general public.

Objective

Newspaper is one of the important sources of information regarding the results of medical research. Health-related articles in newspapers may influence policy makers, consumers of health services and the general public.10–12 The objective of this study is to characterize public health problems related to the use of asbestos in Japan, where asbestos was not fully banned in the 1980s and 1990s. We conducted text mining of newspapers containing articles about social problems to explore the asbestos panic retrospectively.13,14

Subjects and methods

The Asahi Shimbun Newspaper

We reviewed headlines in the Japanese Asahi Shimbun Newspaper with the words either ‘asbestos’ or ‘ishiwata’, which is equivalent for asbestos in Japanese. The databases for the research were CD-ASAX and Kikuzo DNA for libraries. Both were provided by the Asahi Shimbun Company. CD-ASAX contained only headlines and covered the period from 1945 to 1995. In the Kikuzo DNA for libraries, headlines were reviewed from 1996 to 2005, and the morning, evening and late editions of the Tokyo paper were researched. The Asahi Shimbun Newspaper, which has a circulation of 8 000 000, is among the most popular newspapers in the country. The research included text mining with network analysis of the headlines related to asbestos.

Text mining with network analysis

We made a word list of nouns, verbs and adjectives from the headlines using the Japanese language morphological analysis program Cha-sen 2.0 (Nara Institute of Science and Technology) and constructed a matrix correspondent with a headline. The square matrix included word lists in a row and column and indicated the occurrence of a word and the simultaneous occurrence of two words. We summed the matrixes and constructed an annual matrix.

We analyzed the annual matrix using the multidimensional scaling (MDS) PROXSCAL of SPSS 14.0 (SPSS Inc.) and the network analysis of UCINET for Windows, Version 6 (Analytic Technologies Inc.). MDS is a statistical technique for data visualization. We used a matrix of word–word similarities and plotted items according to their similarities in a low-dimensional space. Network analysis is the analysis of networks through graph theory, which studies the properties of graphs designed as a set of nodes connected by ties. Network analysis includes descriptions of structures, such as small-world networks or scale-free networks, and social network analysis, which is the analysis of a set of people or groups of people with some pattern of contacts or interactions between them.15–18 Network analysis is applied in the areas of not only society, internet and nervous system but also languages.19,20 We considered the words as the nodes and the simultaneous occurrence of two words as the ties. We defined some values and indexes as follows:21

(i) The number of headlines (N) represents the total dealing with a specific subject.

(ii) The node-weighted value (NWV) is the number of occurrences of a word.

(iii) The tie-weighted value (TWV) is the number of simultaneous occurrences of two words.
The node-weighted index (NWI) is the percentage of occurrences of a word and is calculated by NWV/N.

The tie-weighted index (TWI) is the percentage of simultaneous occurrences of two words and is calculated by TWV/N.

This methodology is one of the methods for clear visualization. Using the methodology made it possible to structuralize text data, browse it and visualize the characteristics of it to extract new insights.

**Results**

**The Asahi Shimbun Newspaper**

Thirty-six headlines from the *Asahi Shimbun* Newspaper contained the word ‘asbestos’ in 1987 and 293 headlines included it in 2005. The years 1987 and 2005 were selected because the graph indicated that rapid changes had taken place (Fig. 1). The headlines for those years were analyzed through text mining with network analysis.

In 1987, the words ‘pollution’ (NWI = 40%—means that 40% of the headlines included this word), ‘removal’ (31%), ‘campaign’ (29%), ‘expulsion’ (26%), ‘school’ (23%), ‘primary school’ (11%) and ‘junior high school’ (11%) had high NWIs (Table 1). The public organizations and the percentages of times that they were mentioned are as follows: Ministry of Health and Welfare (MHW) (11%), Ministry of Education (MOE) (6%) and the Environment Agency (6%). The combination of words ‘campaign’ and ‘expulsion’ (TWI = 26%—means that 26% of the headlines included both words) and ‘removal’ and ‘campaign’ (14%) were the ties with a high TWI (Table 1).

Figure 2 shows the results of text mining with network analysis for 1987. The nodes for ‘asbestos’ and/or ‘ishiwata’ were deleted because the subject necessarily included either word. This method has the flexibility for visualization. We can change the value of cut-off points of TWV and TWI, compare various versions of cut-off points and find out a version that is no more complicated to interpret. Figure 2A (TWV ≥ 1) could be too complicated to interpret. Figure 2B (TWV ≥ 2) suggested mainly two parts. One part dealt with the campaign to remove asbestos from the public schools. The other dealt with asbestos pollution. Figure 2C (TWV ≥ 3) included only little information.

For 2005, ‘hazard’ (NWI = 31%), ‘person’ (16%), ‘death’ (13%) and ‘workman’s compensation’ (8%) were the words and combinations of words with high NWIs. The public offices and the percentages of times that they appear are listed in Table 1 as follows: the MHLW (6%) and the MOE (4%). The combination of words ‘person’ and ‘death’ (TWI = 9%), ‘harm’ and ‘person’ (8%) and ‘harm’ and ‘death’ (6%) are the ties with high TWIs (Table 1).

Figure 3 shows the results of text mining with network analysis for 2005. We changed the value of cut-off points of TWV and TWI, then we agreed that the version of TWV ≥ 3 (TWI = 3/293 ≈ 1.0%) was too complicated. The version of TWV ≥ 4 suggested more clearly that the main part focused on the area including ‘harm’, ‘person’ and ‘death’.

**Discussion**

**Main findings of this study**

For 1987, the words related with school and asbestos removal have a high NWI. The words MHW and MOE are mentioned, however, the Ministry of Labor (MOL) and the Environmental Agency are not. Most of the reports dealt with asbestos removal and asbestos pollution, and little was written about the potential health hazards related to asbestos exposure. The primary focus was on asbestos removal from the schools, whereas other risks got less attention in the newspaper. Unlike Europe and the United States, Japan continued to import asbestos into the late 1980s (Table 1), which could imply that the efforts to prevent sickness and death among workers and citizens were insufficient.

On the contrary, in 2005, words related with health risks, such as ‘harm’ and ‘death’, had high NWIs, and the newspaper mainly focused on health hazards to the general public represented ‘residents’ (NWI = 6%) in Fig. 3. The names of the public agencies were changed in 2001 as a result of a government reorganization. ‘MOE, Culture, Sports, Science and Technology’ (MEXT), which is old ‘MOE’, is not mentioned unlike in 1987; instead ‘MHLW’, which is old ‘MHW’ and old ‘MOL’, is done.

By applying the method, this study shows the quantitative difference between the contents of the newspaper in 1987 and 2005. The 1987 newspaper focused on asbestos removal from public schools, whereas the 2005 newspaper mainly focused on the health hazards associated with asbestos exposure in 2005.

**What is already known on this topic**

The risks associated with the use of asbestos are well established.\(^1\text{–}^4\) A characteristic of asbestos-linked disease is that there is a long latency period; a latency period of lung cancer is at least 10–20 years.\(^5\text{–}^6\) The risks to general public in non-occupational environments were pointed out more than 15 years ago in Japan.\(^22\text{,}^23\) Reports that asbestos had been used in the construction of schools were widespread in newspapers.
Table 1 Node-weighted index and tie-weighted index

<table>
<thead>
<tr>
<th>1987</th>
<th>Pollution</th>
<th>Removal</th>
<th>Campaign</th>
<th>Expulsion</th>
<th>School</th>
<th>Measure</th>
<th>Junior high school</th>
<th>Inquest</th>
<th>Tokyo metropolitan</th>
<th>MHW</th>
<th>Primary school</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pollution</td>
<td>40%*</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>–</td>
<td>11%</td>
<td>6%</td>
<td>9%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Removal</td>
<td>31%*</td>
<td>14%</td>
<td>11%</td>
<td>9%</td>
<td>–</td>
<td>3%</td>
<td>–</td>
<td>3%</td>
<td>–</td>
<td>11%</td>
<td>–</td>
</tr>
<tr>
<td>Campaign</td>
<td>29%*</td>
<td>26%</td>
<td>11%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>–</td>
<td>6%</td>
<td>–</td>
<td>6%</td>
</tr>
<tr>
<td>Expulsion</td>
<td>26%*</td>
<td>11%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>–</td>
<td>6%</td>
<td>–</td>
<td>6%</td>
</tr>
<tr>
<td>School</td>
<td>23%*</td>
<td>11%</td>
<td>–</td>
<td>3%</td>
<td>9%</td>
<td>3%</td>
<td>3%</td>
<td>9%</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Measure</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3%</td>
<td>–</td>
<td>–</td>
<td>3%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Junior high school</td>
<td>11%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>9%</td>
<td>9%</td>
<td>–</td>
<td>9%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Inquest</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Tokyo metropolitan</td>
<td>11%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>11%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>MHW</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>11%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Primary school</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>11%</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2005</th>
<th>Harm</th>
<th>Person</th>
<th>Death</th>
<th>Measure</th>
<th>Inquest</th>
<th>Usage</th>
<th>Not</th>
<th>Workman’s compensation</th>
<th>MHLW</th>
<th>Year</th>
<th>Mesothelioma</th>
<th>Residents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harm</td>
<td>31%*</td>
<td>8%</td>
<td>6%</td>
<td>2%</td>
<td>3%</td>
<td>1%</td>
<td>3%</td>
<td>2%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>3%</td>
</tr>
<tr>
<td>Person</td>
<td>16%*</td>
<td>9%</td>
<td>–</td>
<td>2%</td>
<td>–</td>
<td>0.3%</td>
<td>1%</td>
<td>1%</td>
<td>–</td>
<td>1%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Death</td>
<td>13%*</td>
<td>0.3%</td>
<td>1%</td>
<td>–</td>
<td>1%</td>
<td>1%</td>
<td>–</td>
<td>2%</td>
<td>3%</td>
<td>3%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Measure</td>
<td>11%*</td>
<td>1%</td>
<td>0.3%</td>
<td>1%</td>
<td>–</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>0.3%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Inquest</td>
<td>10%*</td>
<td>1%</td>
<td>1%</td>
<td>–</td>
<td>1%</td>
<td>1%</td>
<td>–</td>
<td>1%</td>
<td>0.3%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Usage</td>
<td>9%*</td>
<td>–</td>
<td>1%</td>
<td>–</td>
<td>1%</td>
<td>1%</td>
<td>–</td>
<td>1%</td>
<td>0.3%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Not</td>
<td>8%*</td>
<td>–</td>
<td>0.3%</td>
<td>0.3%</td>
<td>0.3%</td>
<td>0.3%</td>
<td>0.3%</td>
<td>1%</td>
<td>–</td>
<td>0.3%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Workman’s compensation</td>
<td>8%*</td>
<td>1%</td>
<td>–</td>
<td>0.3%</td>
<td>1%</td>
<td>0.3%</td>
<td>1%</td>
<td>1%</td>
<td>0.3%</td>
<td>1%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>MHLW</td>
<td>6%*</td>
<td>6%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>6%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Year</td>
<td>6%*</td>
<td>6%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Mesothelioma</td>
<td>6%*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Residents</td>
<td>6%*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

The node-weighted index (NWI) is the percentage of occurrences of a word. The tie-weighted index (TWI) is the percentage of simultaneous occurrences of two words. For example, in 1987, the NWI of ‘pollution’ was 40%, which means that 40% of the headlines contained the word ‘pollution’. The TWI of ‘pollution’ and ‘removal’ was 3%, which means that 3% of the headlines contained the words ‘pollution’ and ‘removal’ in combination. Panel of 1987 shows only the words that have NWI ≥ 11.0% in descending order. Panel of 2005 shows only the words that have NWI ≥ 6.0% in descending order.

*NWI.
Newspaper is one of the important sources of information regarding the results of medical research. Health-related articles in newspapers may influence policy makers, consumers of health services and the general public.10–12

Japanese people were aware of the problem of airborne asbestos use in the late 1980s; however, it seems that a research comparing the quantitative difference between the contents of newspaper in 1987 and 2005 is absent.

What this study adds
This study adds three points.

First, this research shows the difference of the headlines of 1987 and 2005. That is, the 1987 newspaper focused on asbestos removal from public schools, whereas the 2005 newspaper mainly focused on the health hazards associated with asbestos exposure in 2005. From this difference in newspaper headlines, it may be possible to extract the following insight.

There were only few headlines between 1990 and 2004. Many people may have thought the asbestos issue had been resolved, and the risks to the general public were ignored until 2005. After several deaths related to asbestos exposure


**Fig. 2** Network analysis of 1987. The size of nodes is in proportion to a node-weighted value (NWV), which is the occurrence of a word. The thickness of the lines is in proportion to the tie-weighted value (TWV), which is the simultaneous occurrence of two words. The distance of nodes corresponds to similarities by multidimensional scaling (MDS). The node-weighted index (NWI) is the percentage of occurrences of a word and is calculated by NWV/N; the number of headlines (N) represents the total dealing with a specific subject. The tie-weighted index (TWI) is the percentage of simultaneous occurrences of two words and is calculated by TWV/N. This method has the flexibility for visualization by changing the value of cut-off point. This figure shows three optional versions: (A) TWV ≥ 1, (B) TWV ≥ 2 and (C) TWV ≥ 3 (the distance of nodes in (C) is modified because of space limitations). ‘MHW’ is the Ministry of Health and Welfare. ‘MOE’ is the Ministry of Education.
among citizens were reported in 2005, the press began to emphasize the health risks associated with asbestos use, and the government announced plans to implement a total ban by 2008. The lack of an in-depth discussion on this serious issue led to 20 years of silence during which many people were misled into believing that the issue had been resolved. These 20 years of silence ultimately turned out to be a time bomb that exploded when increased asbestos exposure finally resulted in fatalities among the general public in 2005. The consequences of ignoring the problem for 20 years,
which, coincidentally, is compatible with almost the latency period for developing neoplasm, have resulted in widespread exposure to asbestos.

This research does not try to draw causal effect between newspaper and general public way of thinking. It is hard to determine the influence of newspaper information on people’s way of thinking, as the press is not the only source of public health information. However, the clear difference between Fig. 2 and Fig. 3 could imply the above-mentioned insight; the consequences of ignoring the problem for 20 years is compatible with almost the latency period for developing neoplasm, and a supposedly resolved issue turned out to be a time bomb, that is, fatalities among the general public in 2005.

Second, we applied a novel approach of the text mining with network analysis to quantity and characterize the headlines. This study shows one of the methods for visualization of text data. We can reveal the structure of the text data source and extract insights. Additionally, this method has flexibility for visualization as we try to change the values of cut-off points of NWI and TWI. This flexibility enables an exploratory analysis.

Finally, this research methodology suggests that analyzing text data by this method can serve future surveillance and efficient use of epidemiological knowledge. As newspaper is also considered as good sources of information for issues of public health,24 the analysis of newspapers could imply the social concerns of the general public. This research visually described an aspect of the characteristics of the Japan asbestos problem; yet, this methodology may help reveal the characteristics of other various social problems, such as AIDS, smoking, bovine spongiform encephalopathy (BSE), hepatitis B and dioxin.

Limitations of this study
This study, to the best of our knowledge, is the first effort of text mining with network analysis in an effort to
explore the history of the development of asbestos-related health problems in Japan leading to the acute awareness and concern in 2005. However, this study has some limitations.

Newspaper is not the only source of public health information. However, as this method requires electronic text data, newspaper was chosen as the research subject. Moreover, this study checked the headlines of only the *Asahi Shimbun* Newspaper because no other computerized databases have articles older than 50 years.

There are differences in overtones between English and Japanese; for example, ‘Yen’ (Fig. 2), a multisense word in Japanese, means the unit of Japanese money in English, but it has also another meaning, namely, the adjective, ‘round’. To overcome such a limitation, we devoted time to check the exact meaning of words in all headlines.

**Acknowledgements**

We thank Professor Minoru Yabuuchi, the director of the Graduate School of Psychology at Kyoto Notre Dame University, for his valuable suggestions and Michal Nissim for her assistance with the manuscript preparation.

**Competing interests**

None declared.

**References**