The incidence of esophageal squamous cell carcinoma (ESCC), which is the eighth most common malignancy worldwide (Stoner and Gupta 2001), is highest in China and its geographic variation in incidence is striking. In northern China, the southern part of Hebei province, the eastern part of Shanxi province, and the northern part of Henan province have a particularly high incidence of ESCC (Yang 1980; Qiao et al. 2001; Gao et al. 2009; Zhang et al. 2011; Song et al. 2011). Shexian county, which is located in the southern Taihang Mountains, has one of the highest incidences of ESCC in the world, with a rate of 210.5-325.8/10,000 in men and 117.5-185.7/10,000 in women (Zhang et al. 2011). These rates are significantly higher than the Chinese national average of 17.0/10,000 (Wu et al. 2006).

Environmental, biological, and genetic factors have all been implicated in the pathogenesis of esophageal cancer. The factors include low consumption of fresh fruit and green vegetables, insufficient micronutrient intake, nitrosamine, smoking, and alcohol consumption (Stoner and Gupta 2001). In addition, some studies have shown that environmental factors such as drought, high wind speed, normalized difference vegetation index, and elevation are associated with esophageal cancer incidence (Kmet and Mahboubi 1972; Wu and Li 2007; Wu et al. 2008). Drinking water was found to be an important putative environmental factor of the high ESCC incidence in the Linzhou area, which is near the southern part of Shexian county (Han et al. 2007). Consequently, researchers have suspected that drinking water plays the same important role in ESCC pathogenesis in the Shexian area. However, few studies have examined the association between chemical compounds in drinking water and the incidence of ESCC in Shexian. Therefore, the Department of Medical Informatics of Niigata University and the Fourth Affiliated Hospital of Hebei Medical University started a cooperative research
project to investigate such associations in southern Hebei Province. In this paper, we attempt to show the relationship between each risk factor and the incidence of ESCC in the study area, as these findings have the potential to help local governments establish effective cancer prevention policies.

In the past, researchers used paper maps to evaluate and describe cancer distribution, but such maps have limitations. For example, it is difficult to combine layers, and doing so is costly and time-consuming. Recently, geographic mapping has been increasingly used for epidemiology studies, particularly after the development of geographic information system (GIS) technologies that enable data analysis by means of spatial strategies and more comprehensive visual representation of information. GIS is commonly used in environmental epidemiology to assess exposure to pollutants in air (McEntte and Ogneva-Himmelberger 2008), water (Wycisk et al. 2003; Paliwal et al. 2007), and soil (Jarup 2004).

However, only a few reports have posited a relationship between nitrogen compounds in drinking water and the incidence of ESCC in regions with a high incidence of ESCC. This study investigates the association between the incidence of ESCC and the concentration of nitrogen compounds in drinking water of Shexian, China, an area of high ESCC incidence. Spatial distribution of the concentration of nitrogen compounds in the area should increase the understanding of their role in ESCC pathogenesis and lead to the postulation of hypotheses regarding the mechanisms of each compound on the high incidence of ESCC.

Materials and Methods

Fig. 1 shows the location of Shexian county in Hebei province. The period of epidemiological survey was from January to December 2010. We surveyed 52 villages in three areas (Lutou, Gucheng, and Hezhang) of Shexian county. Of the villages, two in Lutou were so close to each other that it was not possible to distinguish a geographic boundary, so both villages were considered a single village. Another four villages share a single well and are within 500 m of each other in the Guxing area, so these villages were also considered a single village. Finally, we collected case data of the 48 villages, which are distributed across three districts. Patient data were obtained from the Shexian Cancer Registry, a population-based registry established in 1974, which is used to register local cancer patients and residents. All esophageal cancer patients were coded according to the International Classification of Diseases. There were 19 cancer patients without ESCC that were excluded. As a result, there were 661 adults with ESCC and 54,055 non-cancer subjects. We then extracted the age-adjusted annual incidence rates of ESCC in 2010.

Water samples were extracted from wells, rivers, and cisterns. In the study area, wells serve all residents in a village, which normally owns one or two wells, although some large villages own four or five wells. Some families obtain water directly from the Qingzhanghe or Zhuozhanghe rivers because their houses are nearer to these rivers than to wells. In all cases, water was sampled from the locations where villagers obtained their daily drinking water. Some villages were located at a high elevation (> 700 m) and were therefore far from surface water. It is difficult and expensive to sink wells at high elevations. Therefore, drinking water cannot be obtained from wells or rivers, so such villages build a large cistern to store rain water as drinking water for residents. The altitude and global positioning system (GPS) coordinates of the water sampling locations, as well as depth for wells, were recorded for each source.

Domestic garbage and agricultural fertilizer were considered as two possible sources of nitrogen contamination in areas of Hebei province with high ESCC incidence (Zhang et al. 2003). We surveyed 86 water sampling points from 60 wells, 10 cisterns, and 16 water points along the Qingzhanghe or Zhuozhanghe rivers. Water samples were extracted twice during 2010: during the low water period (March), and during the high water period (September). Concentrations of nitrogen compounds were measured within 24 h after sampling. The average concentration of nitrate nitrogen, nitrite nitrogen, and ammonia nitrogen, compounds which constitute the principal suspected risk factors for ESCC in the study area (Yokokawa et al. 1999), was recorded and used for analysis (Table 1).

Statistical and Spatial Analysis

R software (R 2.13.1) was used for statistical analysis (Crawley 2007). Spearman’s correlation was used to analyze the correlation between variables. Logistic regression analysis was performed to obtain risk factors for incidence of ESCC. The standardized incidence ratio for ESCC was calculated for each village and mapped to digital map. To detect the high incidence of ESCC and the high concentration of nitrogen compounds in the study area, hot-spot analysis was performed using Getis-Ord Gi* cluster analysis in ArcGis spatial software (ArcGis 9.1.3, Redlands, CA). We used a fixed distance band of 2 km for cluster analysis of incidence. All villages with significant Z scores were identified. Negative values indicate villages with 2-km radius clusters of lower than expected cancer incidence rates, and high Z scores indicate villages with higher than expected cancer incidence rates within a 2-km radius. The digital map was generated by ArcGIS (9.3.1, Redlands, CA).

Results

Esophageal squamous cell carcinoma clusters

The standardized incidence ratio for ESCC was calculated for each village. Results show a statistically significant cluster of villages with a high incidence of ESCC in the southwest region. Low-incidence ESCC clusters occurred predominantly in the northern region. A small low-incidence cluster occurred in the southeast region (Fig. 2).

Nitrogen compound clusters

Villages with a high concentration of nitrate nitrogen (NO₃⁻) were distributed in the southwest region of the study. Conversely, large clusters of villages with low concentrations of NO₃⁻ were distributed in the northern region, along the Qingzhanghe and Zhuozhanghe rivers (Fig. 3). Overall, the concentration of nitrite nitrogen (NO₂⁻) was low in study area. Only a small cluster of villages with a high concentration of NO₂⁻ were distributed in the northwest region (Fig. 4). The high concentration of ammonia nitrogen (NH₄⁺) clusters appeared in the northwest region (Fig. 5). In contrast to the distribution of NO₃⁻, a low
The concentration of NH₄ occurred in the southern region.

Correlation between incidence of esophageal squamous cell carcinoma and nitrogen compounds

There was a weak negative correlation between EC incidence and nitrite nitrogen concentration, ammonia nitrogen concentration and elevation (−0.15, −0.39 and −0.50, respectively), and a positive correlation between EC incidence and nitrate nitrogen concentration (0.38) (Table 2).

Logistic regression analysis was employed to estimate the risk of ESCC incidence. Table 3 presents the estimated coefficient and odds ratio for each candidate risk factor. The risk factor of NO₃ was significantly associated with the
incidence of ESCC in the study area, and the odds ratio was 46.29. NO2, NH4, and elevation had no significant association with the incidence of ESCC.

**Discussion**

Biological, genetic, and environmental factors are thought to play a critical role in ESCC. Although many compounds might be involved in the pathogenesis of ESCC, nitrogen compounds are thought to play the major role in this process. Recent study has shown that nitrogen compounds are not risk factors for cancer (Ward et al. 2008). In contrast, the present study showed that ESCC was influenced by nitrogen compounds in drinking water, especially NO3; high ESCC incidence areas were mostly areas with high concentrations of NO3.

The mechanism of how NO3 affects cancer is complicated and is likely indirect. Although NO2 can be converted to nitrosamine in the stomach (McKnight et al. 1999), the concentration of NO2 in nature is too low to pose a problem in this respect (Table 1). Therefore, it is hard to consider environmental NO2 as a risk factor, since even high-incidence ESCC areas do not have high concentrations of NO2 (Fig. 4). Logistic regression analysis confirmed that NO2 was not a significant risk factor. A high concentration of NO3, on the other hand, is a primary risk factor. van Loon et al. (1998) showed that NO3 has no harmful effect, even when intake is high. However, when NO3 is assimilated in the body, it is converted to NO2 by denitrifying bacteria, and NO2 can subsequently be converted to nitrosamine in the stomach (McKnight et al. 1999). Consequently, a high concentration of NO3 will lead to a high incidence of ESCC. Hot-spot analysis and logistic

<table>
<thead>
<tr>
<th>Water source</th>
<th>Sample points</th>
<th>Nitrate nitrogen (mg/L)</th>
<th>Nitrite nitrogen (mg/L)</th>
<th>Ammonia nitrogen (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well</td>
<td>60</td>
<td>8.08 ± 4.39</td>
<td>0.01 ± 0.05</td>
<td>0.12 ± 0.09</td>
</tr>
<tr>
<td>River</td>
<td>16</td>
<td>5.59 ± 2.39</td>
<td>0.01 ± 0.03</td>
<td>0.12 ± 0.07</td>
</tr>
<tr>
<td>Cistern</td>
<td>10</td>
<td>4.55 ± 2.60</td>
<td>0.04 ± 0.05</td>
<td>0.10 ± 0.07</td>
</tr>
</tbody>
</table>

Fig. 2. Spatial clustering of villages by incidence of esophageal squamous cell carcinoma (per 10,000 population).

Fig. 3. Spatial clustering of villages by nitrate concentration.
Association of Nitrogen Compounds with ESCC

Regression implied this association in our study. In particular, hot-spot analysis showed a high concentration of NO₃ in the middle and southern parts of Shexian county; areas with a high incidence of ESCC (Fig. 3).

Table 2. Correlation of esophageal squamous cell carcinoma incidence with concentration of nitrogen compounds and elevation.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Correlation coefficient</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrate nitrogen</td>
<td>0.38</td>
<td>0.01</td>
</tr>
<tr>
<td>Nitrite nitrogen</td>
<td>−0.15</td>
<td>0.31</td>
</tr>
<tr>
<td>Ammonia nitrogen</td>
<td>−0.39</td>
<td>0.01</td>
</tr>
<tr>
<td>Elevation</td>
<td>−0.50</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Table 3. Regression coefficient ($\beta$), standard error ($SE_\beta$), odds ratio (OR), and 95% confidence intervals (CI) for OR by logistic regression analysis.

<table>
<thead>
<tr>
<th>Factors</th>
<th>$\beta$</th>
<th>$SE_\beta$</th>
<th>OR</th>
<th>95% CI for OR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.77</td>
<td>0.32</td>
<td>46.29</td>
<td>3.16 - 667.39</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Nitrate nitrogen</td>
<td>3.84</td>
<td>1.37</td>
<td>0.29</td>
<td>0.05 - 1.68</td>
<td>0.01</td>
</tr>
<tr>
<td>Nitrite nitrogen</td>
<td>−1.24</td>
<td>0.90</td>
<td>0.34</td>
<td>0.06 - 1.84</td>
<td>0.17</td>
</tr>
<tr>
<td>Ammonia nitrogen</td>
<td>−1.09</td>
<td>0.87</td>
<td>0.82</td>
<td>0.15 - 4.59</td>
<td>0.21</td>
</tr>
<tr>
<td>Elevation</td>
<td>−0.19</td>
<td>0.88</td>
<td>0.82</td>
<td>0.15 - 4.59</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Ammonia nitrogen correlated negatively with the incidence of ESCC in this study. In the middle and southern parts of the study area, villagers use rivers and wells as their main water source. Both sources have fresh
water because it is not stored for long periods. In contrast, northern villages use cisterns to store water for long periods because it is difficult to obtain drinking water from rivers and wells. In this case, stored drinking water is at higher risk of contamination with organic matter than flowing water, and the concentration of NH4 is higher in the northern villages than in the middle and southern villages. Despite these findings, the high incidence of ESCC was in the middle and southern parts, while the northern parts had a relatively low incidence of ESCC (Fig. 5). Therefore, in this case, the correlation between NH4 and the incidence of ESCC is negative. A more accurate correlation between ESCC and NH4 should be confirmed using a high number of water samples and comparing them with areas of high and low incidence of ESCC.

The incidence of many cancers is associated with elevation (Krain 1991). Wu et al. (2008) suggested that high altitude might decrease the incidence of EC. Our study confirmed this negative correlation between elevation and ESCC incidence. Spearman’s correlation test showed that the correlation coefficient of elevation and incidence was −0.50 (P < 0.001) (Table 2). Hot-spot analysis showed that high elevated areas have a low incidence of ESCC.

The present study has several limitations. This study focused on drinking water in residential areas, but water sampling was not comprehensive as, for example, water samples were taken only from the middle and lower reaches of rivers making it impossible to determine the source of river pollution. Furthermore, we focused mainly on nitro-gen compounds in drinking water and did not collect information on land use (e.g., agricultural versus industrial). Thus, there is no way of determining whether surface water and groundwater were polluted by industrial or agricultural wastewater. In addition, other known ESCC risk factors, such as biological and genetic factors, were not investigated in the present study. Thus, further studies are needed to address potential risk factors, preferably by using individual level data on health.

In conclusion, we have established an association between nitrogen compounds in drinking water and the incidence of ESCC; consequently, high concentrations of nitrate nitrogen in drinking water may increase the risk of ESCC. Furthermore, residents living in the southwest part of Shexian county were found to carry a particularly high risk of ESCC.

Acknowledgments

We appreciate the participation of all subjects and persons in this study. We would like to express our gratitude to David Kipfer for proofreading on our manuscript.

Conflict of Interest

All authors declare no conflict of interest.

References


