Microbial contamination and environmental factors of drinking water source for households with children under five years old in two South China ethnic groups

Wen Fang Long, Spencer Lloyd, Fan Zhang, Qiao Li, HaiRong Huang and Cong Yi

ABSTRACT

The article aimed to find the causes of microbial contamination of drinking water source, by exploring the influencing factors of water on the incidence of diarrhea in children. Random stratified-cluster sampling was used to study Li and Han communities of rural Lingshui County with regard to water quality and environmental sanitation. Water samples were evaluated for microbial contamination using National Standard of the People's Republic of China and World Health Organization drinking water standards. Data were compared between ethnic groups. Li communities had more contamination than Han. The correlation between the frequency of diarrhea among children and the presence of microorganisms in the water was positive, for both total coliforms and *Escherichia coli* (*E. coli*). Regression analysis showed the substandard risk factors of total coliforms for source water were type of water supply (OR = 3.508) and garbage disposal methods (OR = 2.430). For *E. coli*, risk factors included the source of water supply (OR = 2.417); depth of wells (OR = 0.536) and distance of wells from the cesspit (OR = 0.723). The content of bacterium in drinking water source was high in the rural county. Water from open wells had higher contamination rates than water from centralized systems and tube-well sources. Improvement of water supply and environmental hygiene would decrease diarrheal diseases among children under five.

Key words | China, diarrhea, drinking water, *Escherichia coli*, total coliforms

INTRODUCTION

The annual global mortality of children under five years old is around 8.8–10.6 million (Kosek et al. 2005; Bryce et al. 2005). Diarrhea accounts for eighteen percent of these deaths and, therefore, is one of the top three causes of death in children under five years old in developing countries around the world (Murray et al. 2007). This study evaluated water supply, quality, hygiene and sanitation factors that might contribute to increased gastrointestinal diseases among children under five years old in a less developed county in South China. Lingshui County is one of the poorest counties in Hainan Province, and has the largest Li ethnic minority group in China (>90%). In order to find the influencing factors of drinking water on the incidence of diarrhea in children living in Lingshui County, an investigation was carried out focused on environmental health and microbial contamination of drinking water in the summer of 2012.

MATERIALS AND METHODS

Sample selection

Sample size for one ethnic group is calculated using the following formula (Higgins & Shulman 1994):

\[ N_1 = \frac{p(1-p)(Z_{a/2})^2}{d^2} \]

Wen Fang Long
Fan Zhang
Qiao Li
HaiRong Huang
Cong Yi (corresponding author)
School of Public Health
Hainan Medical University, No. 3 Xueyuan Road, Longhua District, Haikou 571199, Hainan Province, China
E-mail: yi_cong@hotmail.com

Spencer Lloyd
Infectious Disease Fellow, Division of Infectious Diseases,
Department of Internal Medicine, University of Utah School of Medicine, 30 North 1900 East, Room 4B319, Salt Lake City, UT 84132, USA
where $N_1$ is the required sample size, $p$ is proportion of interest, $Z_a$ is 1.96 for a 95% confidence interval and $d$ is the distance, or how close we require our estimate to be to the proportion of interest. This formula yields a sample size of approximately 80. By using a clustering method, we introduce a design effect ($Deff$), which requires an increased sample size. If each rural village has an average of 25 children under five years old then the $Deff$ is calculated by (Bland 2004):

$$Deff = 1 + (m - 1) \times ICC$$

where $m$ equals the cluster size and ICC is the intraclass correlation coefficient. A typical ICC is 0.01–0.02 for human studies (Murray et al. 1994; Murray & Short 1995, 1997). If we let ICC equal 0.02 then $Deff$ equals 1.48 and our final sample size is $N_1 \times Deff$ or $80 \times 1.48 = 118$. The $Deff$ value can vary substantially depending on the indicator investigated and on the context, but according to guidelines published by the Department of International Health at Johns Hopkins University, $Deff$ values encountered in different cluster survey projects range from 1.5 to 2.0 (Sarriot et al. 1999). If we use a $Deff$ value of 2.0 then our sample size increases to 160. Given our uncertainty on the prevalence of childhood diarrhea in Hainan Province, we will increase our sample size per ethnic group, substantially depending on the indicator investigated and on the context, but according to guidelines published by the Department of International Health at Johns Hopkins University, $Deff$ values encountered in different cluster survey projects range from 1.5 to 2.0 (Sarriot et al. 1999). If we use a $Deff$ value of 2.0 then our sample size increases to 160. Given our uncertainty on the prevalence of childhood diarrhea in Hainan Province, we will increase our sample size per ethnic group. Therefore, the total sample size will be 400 children under five years of age.

Two towns of Li ethnic and two towns of Han ethnic were sampled. With 100 families with children under five years old from each town, a total of 400 families were selected as study sample.

**Questionnaire respondents**

The towns, villages, and families were stratified randomly according to the different Li and Han populations. Twenty villages from four towns, two Li towns and two Han towns were sampled to investigate the drinking water quality and environmental sanitation.

**Survey contents**

Questions were focused on general hygiene practices and included types of water source, protection of water source, toilet hygiene, garbage disposal, childhood diarrhea, etc. Trained investigators conducted the household inquiry, filled out the questionnaire and then inspected for validity and reliability.

The procedures were reviewed and approved by the Institutional Review Boards for use of human subjects in research at Hainan Medical University and the University of Utah, and were in accordance with the Declaration of Helsinki as revised in Edinburgh in 2000. All human subjects gave informed consent and the anonymity was preserved.

**Methods for water sampling, data calculating and analysis**

Water samples were collected directly from all the sampled households. Two parallel samples were collected from households with centralized water sources and one sample was collected from households with other types of source water. Collection and analysis of water samples were conducted according to the water standard of China (Standards for Drinking Water Quality GB5749, Standard Inspection Methods for Drinking Water, GB5750-2006, China), total coliforms and *Escherichia coli* (*E. coli*) were detected by Defined Substrate Technology, DST®.

The number of total coliforms and *E. coli* per 100-ml sample was calculated by using the most probable number (MPN) formula.

$$\text{MPN} = \frac{P}{(N \times T)^{1/2}} (\text{GB/T5750} - 2006)$$

where $P = \#$ of positive wells, $N = \text{total volume of all samples}$, and $T = \text{volume of negative wells}$.

The water samples were evaluated according to the criteria of the Chinese (GB5749-2006) and World Health Organization (WHO) Guidelines for Drinking-water Quality (both are 0 cfu/100 mL) (National Standard of the People’s Republic of China 2006; WHO 2011). The questionnaire data were statistically analyzed using the SPSS16.0 software. This included the descriptive analysis for microbial content. The comparison of qualified ratio for bacterial content in the different villages, towns, and ethnic groups was analyzed by the chi-squared test. The Spearman rank correlation analyses between the bacterial content and the environmental health factors were then made, as well as factors influencing the qualified rate of
water for microorganism criteria using logistic regression analysis. Test levels were of $\alpha = 0.05$.

## RESULTS AND DISCUSSION

### General status

Abnormal test results were analyzed and appropriately disposed of if necessary, without creating a significant difference. The data from 340 valid water samples were processed and are representative of the overall data. Classification according to the water supply: centralized water supply accounts for 29.4% (100 samples) and scattered water supply accounts for 70.6% (240 samples). For scattered water supply households: well water is 69.2% (tube well accounting for 17.2% and open wells for 52%), others are pool water and some spring water. Classification according to the depth of well: shallow well water 69.6%, deep well water 4.0%, others are 26.3%. Therefore, the most common type of drinking water supply in rural Lingshui County is scattered shallow well water.

### Microbial contamination of drinking water source

The content of total coliforms in drinking water source is high. The qualified rate of water samples (the content of bacterium determined to be safe according to water quality standard) is 6.7% (16/240) for total coliforms and 54.6% (131/240) for E. coli, respectively (Table 1).

### Comparison of microorganism contamination of drinking water source in different ethnic groups

The content of both total coliforms and E. coli in the Li group are higher than in the Han group ($P < 0.05$, Table 2). For total coliforms, the qualified rates are 0.5% in Li towns and 26.3% in Han towns, respectively; for E. coli the figures are 46.4% (85/183) in Li towns and 82.5% (47/57) in Han towns. The chi-square shows that the differences between the qualified rates are statistically significant (total coliforms: $\chi^2 = 35.666, P = 0.000$; E. coli: $\chi^2 = 24.492, P = 0.000$).

### Analysis of microbial contamination of drinking water with diarrheal and environmental factors

(1) **Correlation analysis of the bacterial of drinking water with diarrhea**

The prevalence rate of diarrheal disease is 23.2% in the previous 2 weeks. There was correlation between the total coliforms of source water and the presence and severity of diarrhea in children in the 2 days prior to the survey (total coliforms: $r = 0.304, P = 0.017$; E. coli: $r = 0.290, P = 0.020$), but not with diarrhea in the previous 2 weeks.

(2) **Logistic regression analysis for drinking water with environmental factors**

To understand the impact of various factors on total coliforms and E. coli in the source water, multivariate logistic regression analyses were performed. Multivariate logistic regression (Table 3) analysis showed that the source of drinking water, the depth of the wells, and the distance from the feces to well are all risk factors for E. coli exceeding the minimum standard. Logistic regression shows that the source of water supply is the most important risk factor that affects whether water microbial indicator is significant qualified or not, OR = 3.033 (total coliforms) and OR = 2.410 (E. coli). Analysis implies that concentrated waste disposal has an important effect on microbial contamination on a broad spectrum for total coliforms ($P < 0.05$), but with less prevalent effect on E. coli ($P > 0.05$).

Furthermore, the types of scattered water sources were categorized among the Li ethnic and show the following: tap water accounts for 2.1%, tube wells account for 2.7%, protected wells account for 56.3%, unprotected wells account for 37.8% and others are 1.0%. Within the Han towns, tap water...
is 13.8%, tube wells are 56.9%, protected wells are 24.1%, unprotected wells are 5.2%. Thus, Han towns mainly use tube wells and centralized water supply, while Li ethnic towns use mainly scattered well water supply. Centralized water supply of Han towns and tube wells account for more than 70.7%, which is distinctly higher than that of the Li minority, which accounts for only 4.8%. The chi-squared test proves that the different water sources produce a significant difference ($\chi^2 = 4.448$, $P = 0.035$). Further analysis of the differences between diverse water sources indicates that the tube well is safer than open well (including protected and unprotected wells) ($P < 0.05$), protected wells are safer than unprotected wells and other water sources, as shown by differences in total coliforms percentages ($P < 0.05$).

This study shows that the microbial contamination of drinking water is very serious in Li residential regions. The content of bacterium of source water is much higher than that of drinking water in most rural districts in China. The microbial contamination of Li towns, both total coliforms and $E. coli$ is more prevalent than in Han towns in rural Lingshui as well. The findings in this study suggest that transforming water sources in Li towns is urgently needed.

They suggest that the depth of wells is an influencing factor for $E. coli$ contamination, the content of $E. coli$ of shallow wells is higher than that of deep wells. By comparing the depth of different types of wells, we found that the number of shallow wells is greater than that of deep wells in the studied communities. Wells should be deepened in order to avoid fecal-borne pollution.

Distance of wells from the cesspit is a risk factor for $E. coli$ contamination ($P < 0.05$). The distance should be maximized whenever possible to decrease microbial contamination from human and animal feces.

This study does not suggest that the well materials play an important role in water quality in the investigated county. Meanwhile, the coliforms value peaked and the water chroma and turbidity were very poor in few wells using bricks as well wall material in town A (Li group). Total coliforms have an origin in the natural environment and feces of warm-blooded animals or humans; however, $E. coli$ merely comes from warm-blooded animals or human fecal pollution. Town A is located in the hilly region, thus there may be another source from the Jungle. In these cases, transformation of wall materials will make an improvement in water quality.

The prevalence rate of diarrheal diseases among children under five years old within 14 days previous to the survey in Li communities is much higher than that of children in western China counties, which were less developed areas in China as well (Wu 2008). It is also higher than the rate in other developing countries (Nasrin et al. 2013). $E. coli$ contamination of

### Table 2

<table>
<thead>
<tr>
<th>Groups</th>
<th>Total coliforms (MPN/100 mL)</th>
<th>E. coli (MPN/100 mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median (MPN/100 mL) Q1–Q3 (MPN/100 mL)</td>
<td>Median (MPN/100 mL) Q1–Q3 (MPN/100 mL)</td>
</tr>
<tr>
<td>Han</td>
<td>4.54 1.41–56.86 0.00 0–0</td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>51.66* 18.71–141.10 1.50* 0–4.54</td>
<td></td>
</tr>
</tbody>
</table>

*Compare with Han ($P < 0.05$); Q1: Lower quartile; Q3: Upper quartile.

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E</th>
<th>Wald</th>
<th>P</th>
<th>OR</th>
<th>95.0% CI for OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E. coli$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source water supply</td>
<td>0.882</td>
<td>0.280</td>
<td>9.897</td>
<td>0.002</td>
<td>2.417</td>
<td>1.395–4.187</td>
</tr>
<tr>
<td>Distance from well to cesspit</td>
<td>−0.325</td>
<td>0.136</td>
<td>5.742</td>
<td>0.017</td>
<td>0.723</td>
<td>0.554–0.943</td>
</tr>
<tr>
<td>Depth of well</td>
<td>−0.625</td>
<td>0.301</td>
<td>4.295</td>
<td>0.038</td>
<td>0.536</td>
<td>0.297–0.967</td>
</tr>
<tr>
<td>Constant</td>
<td>−1.445</td>
<td>1.074</td>
<td>1.811</td>
<td>0.178</td>
<td>0.236</td>
<td></td>
</tr>
<tr>
<td>Total coliforms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source water supply</td>
<td>1.255</td>
<td>0.330</td>
<td>14.491</td>
<td>0.000</td>
<td>3.508</td>
<td>1.838–6.693</td>
</tr>
<tr>
<td>Garbage disposal</td>
<td>0.888</td>
<td>0.296</td>
<td>8.978</td>
<td>0.003</td>
<td>2.430</td>
<td>1.360–4.345</td>
</tr>
<tr>
<td>Constant</td>
<td>−20.556</td>
<td>4.906</td>
<td>17.55</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

B: Regression coefficient value; S.E: standard error; Wald: chi-square value; P: probability value; OR: odds ratio.
drinking water is highly related to childhood diarrhea. In a diarrheal endemic foci in India, more than one-fifth of the targeted sources (21.4%) have been identified harboring *E. coli* (Batabyal et al. 2013). The study in Bamako, Mali, also reveals similar conclusions (Baker et al. 2013). So measures should be taken to decrease *E. coli* contamination and, thus, decrease childhood diarrhea in rural Lingshui communities, especially Li ethnic communities.

**CONCLUSIONS**

In summary, drinking water quality in rural Lingshui County is poor, especially in Li ethnic communities. The primary influencing factors for bacterium contamination of drinking water are the type of water source, depth of wells and distance of wells from the cesspit. Thus, measures of transforming water supply systems and improving well water quality should be taken to control bacterium contamination, thus decrease childhood diarrhea.

**ACKNOWLEDGEMENTS**

The study was supported by the University of Utah School of Medicine, USA; Hainan Medical University, China; the local government of Lingshui County; and the Provincial Natural Science Foundation of Hainan (No. 813200). We would like to thank all the organizers and implementers for their participation in this study. We are also grateful to all the investigators for their contribution to data collection, especially Martine Troy, Samantha Jeppsen and Debbie Dixon.

**REFERENCES**


National Standard of the People’s Republic of China 2006 *Standards for Drinking Water*. Chinese Center for Disease Control and Prevention/Ministry of Health of China (GB5749, GB5750), Beijing, China.

Sarriot, E., Winch, P., Weiss, W. M. & Wagman, J. 1999 *Methodology and Sampling Issues for KPC Surveys*. Johns Hopkins University, School of Public Health, Department of International Health, Baltimore, MD, USA.
