SHORT COMMUNICATION

Using Bayesian belief networks to support health risk assessment for sewer workers

K. F.-R. Liu · C.-W. Chen · Y.-S. Shen

Received: 15 January 2012/Accepted: 2 August 2012/Published online: 23 November 2012 © CEERS, IAU 2012

Abstract The sanitary sewerage connection rate is an important indicator of advanced cities. Following the construction of sanitary sewerages, the maintenance and management systems are required for keeping pipelines and facilities functioning well. These maintenance tasks often require sewer workers to enter the manholes and the pipelines, which are confined spaces short of natural ventilation and have the potential for hazardous substances to be present. Working in sewers could be easily exposed to a risk of adverse health effects. This paper proposes the use of Bayesian belief networks as a higher level of noncarcinogenic health risk assessment of sewer workers. On the basis of the epidemiological studies, the actual hospital attendance records and expert experiences, the Bayesian belief networks is capable of capturing the probabilistic relationships between the hazardous substances in sewers and their adverse health effects, and accordingly inferring the morbidity and mortality of the adverse health effects. The provision of the morbidity and mortality rates of the related diseases is more informative and can alleviate the drawbacks of conventional methods.

K. F.-R. Liu (🖂)

Department of Safety, Health and Environmental Engineering, Ming Chi University of Technology, New Taipei 24301, Taiwan, ROC e-mail: kevinliu@mail.mcut.edu.tw

C.-W. Chen Institute of Maritime Information and Technology, National Kaohsiung Marine University, Kaohsiung 80543, Taiwan, ROC

Y.-S. Shen

Holistic Education Center, Mackay Medical College, New Taipei 25245, Taiwan, ROC **Keywords** Sanitary sewerage · Morbidity · Mortality · Hazard quotient · Target organ-specific hazard index

Introduction

In terms of the length, 25.8 % of rivers in Taiwan are moderately or severely polluted and discharging of domestic sewage into rivers is one of the main causes. To reduce the river pollution and improve residents' living environment and health, the fundamental solution is to establish sanitary sewerages. A sanitary sewerage is a separate underground carriage system specifically meant for transporting sewage from houses and commercial buildings to treatment or disposal. Indeed, it is widely recognized that the sanitary sewerage connection rate becomes an important indicator of advanced cities. Although the sewerage connection rates of the major cities in Taiwan still lag far behind those in the cities in other developed countries, they are progressing rapidly. Taipei City, a metropolis in northern Taiwan, reached the sewerage connection rate of 100 % in 2010; In Kaohsiung City, a metropolis in the south, the sewerage connection rate was enhanced to 60.89 % in 2010 from 6.5 % of 1989.

Following the construction of sanitary sewerages, the maintenance and management systems are needed to keep pipelines and facilities functioning well. There are three primary tasks in the maintenance and management systems: inspecting pipelines and facilities, investigating the integrity of the sewerage, and dredging obstructions in pipelines. These maintenance tasks often require sewer workers to enter the manholes and the pipelines. However, since the underground sewers are confined spaces short of natural ventilation and the sewage generates hazardous



substances such as toluene, trichloroethylene, trichloromethane, tetrachloroethylene, carbon monoxide, xylene and hydrogen sulfide, working in sewers could easily lead to a risk of adverse health effects. Therefore, some researchers devoted themselves to study the health risk assessment (HRA) of sewer workers (Lai et al. 2004; Yeh et al. 2011).

HRA is still an important issue in environmental studies (Batayneh 2012; Shyam and Kalwania 2012; Tahir et al. 2012; Xue et al. 2012; Liu et al. 2012). HRA is the process for estimating the nature and probability of adverse health effects in humans who may be exposed to hazardous substances. Its four operational stages include hazard identification, exposure assessment, dose-response assessment, and risk characterization (California EPA 2001). Hazard identification aims to recognize any potential health problem that a substance can cause; exposure assessment determines the amount, duration, and pattern of exposure to the substance; dose-response assessment estimates how much of the substance it would take to cause varying degrees of adverse health effects; and risk characterization interprets the risk for the substance to cause cancer or other illnesses. Conventionally, the risk characterization of noncarcinogenic substances can be evaluated by the hazard quotient (HQ), the ratio of the intake of a hazardous substance to its reference dose. Multiple hazardous substances may affect the same organ (or organ system) causing joint effect, and hence the target organ-specific hazard index (TOSHI) sums the HQ scores of multiple substances that have joint effect on a specific organ (US EPA 2001). Several problems arise from using HQ and TOSHI. First, the HQ score for non-critical organs is overestimated because it is derived from the reference dose of the most critical organ with the lowest no-observed-adverse-effect level (NOAEL). Second, reference doses are usually derived from animal studies and it is difficult to precisely manage the uncertainty in animal-to-human extrapolations. Third, as it sums all HQ scores as a single score, TOSHI may be inappropriate for evaluating the joint effect on a specific organ because it is valid only if a common toxic mechanism exists. Fourth, it is unlikely for a TOSHI score below 1.0 to result in adverse noncancer health effects over a lifetime of exposure, whereas a score above 1.0 does not necessarily suggest a likelihood of adverse effects (US EPA 2001). Compared with the probability, risk characterization of the TOSHI is less clear or precise.

To solve these problems, this paper proposes the use of Bayesian belief networks (BBN) (Pearl 1988) as a higher level of HRA, denoted as the BBN-HRA, for noncarcinogenic HRA of sewer workers. The BBN is a directed acyclic graph with nodes denoting a set of random variables as nodes and arrows indicating their probabilistic cause-effect dependencies. There are abundant previous studies on the application of the BBN to environmental issues . They can be roughly divided into four groups: prediction, evaluation, diagnosis, and classification. For example of prediction cases, Liao et al. (2010) used BBN to predict the rate of human neural tube defects by considering the number of doctors, the use of pesticides and fertilizer, the production of vegetable and fruit, per-capita net incomes, elevation, NDVI, road and fault buffer, influence of coal mines, and distances to the nearest factory. An example study of evaluation was the work of Ticehurst et al. (2010) who used the BBN to complement conventional analyses for exploring landholder management of native vegetation. Dawsey et al. (2006) proposed a diagnosis case, which used the BBN to integrate sensor data with other validating evidence of contamination scenarios and then to identify the most probable contamination release nodes in drinking water distribution systems. A classic example can be found in the study of Newton (2010) who used the BBN to produce Red List classifications of threatened species for taxa in situations where the input data are uncertain.

The BBN is capable of capturing the probabilistic relationships between the hazardous substances in sewers and their (critical and non-critical) adverse health effects and accordingly inferring the morbidity and mortality of the adverse health effects if pollution concentrations are given. The provision of the morbidity and mortality rates of the related diseases is more informative and can alleviate the uncertainty of the TOSHI.

The organization of this paper is as follows: the study area is introduced in Sect. 2. In Sect. 2, a brief description of the BBN, the development procedure of the BBN-HRA for the case of sewer workers is elaborated. In Sect. 3, the study area is introduced and its analysis results are then presented to illustrate this model. Finally, conclusions and future work are included in Sect. 4. This research has been carried out in the Department of Safety, Health and Environmental Engineering, Ming Chi University of Technology (Taipei; Taiwan), from January to August 2011.

Materials and methods

Brief introduction of BBN

A Bayesian belief network consists of a directed acyclic graph and an associated computational structure. In the graph, the nodes represent random variables (X_i) with several possible states while arrows connect pairs of nodes to display their probabilistic cause–effect relationships. Each node with parents has a conditional probability distribution table (CPT) that quantifies the uncertain effects the parents have on the node, and those nodes without



Fig. 1 a Example for explaining mathematical basis of BBN; b a given probability distribution {0.8, 0.2} over {high, low} of T and the associated inferred probabilities of other variables; \mathbf{c} a given probability of 0.01 of A and the associated inferred probabilities of other variables

parent has a probability distribution over all possible states. These probabilities are evaluated from historical data, expert judgment, or their combination. For example, a root node (T) can cause other nodes (A) and (B), and further leading to a node (C), as shown in Fig. 1a. T has two possible states {high (h), low (l)} and its associated probability distribution is {8.2 %, 91.8 %}. The possible states of the other three variables are true (t) or false (f) and their respective CPTs, as shown in Fig. 1a.

In the computational structure of the BBN, the joint probability distribution (JPD) of a set of related variables is inferred from the observation. The JPD over variables $\{X_1, X_2, \ldots, X_n\}$ can be formally represented as the product of these CPDs:

$$P(X_1, X_2, ..., X_n) = \prod_i P(X_i | X_{i-1}, ..., X_1)$$
(1)

In Fig. 1 of the Total Suspended Particles (TSP) example, the JPD over variables $\{T, A, B, C\}$ is

$$P(T, A, B, C) = P(C|A, B, T) \times P(B|A, T) \times P(A|T)$$

$$\times P(T)$$
(2)

In Eq. (1), the calculation complexity can be largely reduced if the conditional independence between some variables can be determined. In the graph theory, the conditional independence can be identified through the concept of "d-separation." Assume that X, Y, and E are three sets of nodes. If every undirected path from a node in X to a node in Y is d-separated by E, then X and Y are conditionally independent given E. A set of nodes E d-separates two sets of nodes X and Y if every undirected path from a node in X to a node in Y is blocked given E. A path is blocked given a set of nodes E if there is a node Z on the path for which one of the following three conditions holds: Z is in E and Z has one arrow on the path leading in and another arrow on the path leading out (serial connection); Z is in E and Z has two paths with arrows leading out (diverging connection); neither Z nor any descendant of Z is in E, and arrows on both paths lead into Z (converging connection). In a case of three variables with serial connection $(X_1 \rightarrow X_2 \rightarrow X_3)$ or another case of three variables with diverging connection $(X_1 \leftarrow X_2 \rightarrow X_3), X_1$ and X_3 are conditionally independent given X_2 because X_2 "d-separates" X_1 and X_3 , which permits $P(X_3|X_2, X_1) =$ $P(X_3|X_2)$. However, in a case of three variables with converging connection $(X_1 \rightarrow X_2 \leftarrow X_3)$, X_1 and X_3 are conditionally dependent given X_2 , which makes $P(X_3|X_2,$ $X_1 \neq P(X_3|X_2)$. The identification of conditional independence between some variables can enormously simplify Eq. (1) as

$$P(X_1, X_2, \dots, X_n) = \prod_i P(X_i | \text{parents of } X_i)$$
(3)

The rule of total probability can help derive $P(X_i = x_i)$ by marginalization, i.e. summing out over "irrelevant" variables:



where x_i is the state of variable X_i .

Take again the example shown in Fig. 1b. *T* and *C* are conditionally independent given $\{A, B\}$; *A* and *B* are conditionally independent given *T*; thereby the JPD over variables $\{T, A, B, C\}$ (Eq. 2) becomes

$$P(T, A, B, C) = P(C|A, B) \times P(B|T) \times P(A|T) \times P(T)$$
(5)

Assume that the probability distribution over {high, low} of *T* is predicted as {80 %, 20 %}; therefore, P(T = h, A = t, B = t, C = t) is computed by $P(C = t|A = t, B = t) \times P(B = t|T = h) \times P(A = t|T = h) \times P(T = h) = 0.91729 \times 0.1673 \times 0.09874 \times 0.8 = 0.01212 = 1.212$ %. Similarly, other JPDs can be calculated, as shown in Table 1. Finally, P(A = t), P(B = t) and P(C = t) can be computed by Eq. (4) as 9.25, 15.73, and 7.88 %, respectively, as shown in Fig. 1b.

The previous example is one type of inference whose evidences are given to the root node of a BBN and the probabilities of the other nodes are then derived. In another type of inference, evidences are provided in the intermediate nodes of a BBN and the probabilities of the other nodes can be computed through Eq. (3) and Bayes' rule. If the probability of *A* is anticipated to be improved from 7.03 to 1.00 %, Eq. (5) should be rewritten as

Table 1	Calculations	of joint	probability	distribution	in I	Fig.	11
							_

State of variable			$P(T,A,B,C) = P(C A,B) \times P(B T) \times P(A T) \times P(T)$	
Т	Α	В	С	
h	t	t	t	$0.01212 = 0.91729 \times 0.1673 \times 0.09874 \times 0.8$
h	t	t	f	$0.00109 = 0.08271 \times 0.1673 \times 0.09874 \times 0.8$
h	t	f	t	$0.01568 = 0.23840 \times 0.8327 \times 0.09874 \times 0.8$
h	f	t	t	$0.03281 = 0.27204 \times 0.1673 \times 0.90126 \times 0.8$
h	t	f	f	$0.05010 = 0.76160 \times 0.8327 \times 0.09874 \times 0.8$
h	f	t	f	$0.08781 = 0.72796 \times 0.1673 \times 0.90126 \times 0.8$
h	f	f	t	$0.00621 = 0.01034 \times 0.8327 \times 0.90126 \times 0.8$
h	f	f	f	$0.59418 = 0.98966 \times 0.8327 \times 0.90126 \times 0.8$
1	t	t	t	$0.00146 = 0.91729 \times 0.1171 \times 0.06773 \times 0.2$
1	t	t	f	$0.00013 = 0.08271 \times 0.1171 \times 0.06773 \times 0.2$
1	t	f	t	$0.00285 = 0.23840 \times 0.8829 \times 0.06773 \times 0.2$
1	f	t	t	$0.00594 = 0.27204 \times 0.1171 \times 0.93227 \times 0.2$
1	t	f	f	$0.00911 = 0.76160 \times 0.8829 \times 0.06773 \times 0.2$
1	f	t	f	$0.01589 = 0.72796 \times 0.1171 \times 0.93227 \times 0.2$
1	f	f	t	$0.00170 = 0.01034 \times 0.8829 \times 0.93227 \times 0.2$
1	f	f	f	$0.16292 = 0.98966 \times 0.8829 \times 0.93227 \times 0.2$

h high, l low, t true, f false



$$P(T, A, B, C) = P(C|A, B) \times P(B|T) \times P(T|A) \times P(A)$$
(6)

Moreover, P(T = h|A = t) can be derived by Bayes' rule:

$$P(T = h|A = t)$$

$$= \frac{P(A = t|T = h) \times P(T = h)}{P(A = t|T = h) \times P(T = h) + P(A = t|T = l) \times P(T = l)}$$

$$= \frac{0.009874 \times 0.082}{0.009874 \times 0.082 + 0.006773 \times 0.918}$$

$$= 0.11522$$
(7)

where P(T) is the prior probability of *T* and P(T|A) is the posterior probability given *A*. Similarly, P(T = l|A = t), P(T = h|A = f) and P(T = l|A = f) are obtained as 88.48, 79.49, and 92.05 %, respectively. Finally, P(T = t), P(B = t) and P(C = t) can be computed by Eq. (4) as 7.98, 12.11, and 4.48 %, respectively, as shown in Fig. 1c.

Identification of key factors and their causal relationships

The primary hazardous substances which can cause non-carcinogenic effects in sanitary sewerage are toluene, trichloroethylene, trichloromethane, tetrachloroethylene, carbon monoxide, xylene, and hydrogen sulfide. According to related research (Lin 2005), the non-carcinogenic effects of the seven hazardous substances are summarized in Table 2. Toluene can lead to central nervous system diseases, dermatitis, kidney diseases, and cardiovascular diseases; trichloroethylene induces dermatitis, liver

 Table 2 Hazardous substances in sanitary sewerage and their noncarcinogenic effects

Air pollutant	Adverse health effects				
Toluene	Central nervous system diseases Dermatitis Kidney diseases				
	Cardiovascular diseases				
Trichloroethylene	Dermatitis				
	Liver complaint				
	Cardiovascular diseases				
Trichloromethane	Central nervous system diseases				
	Kidney diseases				
	Liver complaint				
Tetrachloroethylene	Kidney diseases				
	Liver complaint				
Carbon monoxide	Cardiovascular diseases				
Xylene	Cardiovascular diseases				
	Respiratory system diseases				
Hydrogen sulfide	Respiratory system diseases				





complaint, and cardiovascular diseases; trichloromethane can cause central nervous system diseases, kidney diseases, and liver complaint; tetrachloroethylene induces kidney diseases and liver complaint; carbon monoxide can lead to cardiovascular diseases; xylene can cause cardiovascular diseases and respiratory system diseases; and exposure to hydrogen sulfide can result in respiratory system diseases. These hazardous substances and the induced diseases are the nodes in the BBN, as shown in Fig. 2.

Development of Cpts

Discretization of continuous variables

The seven hazardous substances are essentially continuous variables. However, most BBN software cannot deal with continuous variables so that the solution is to discretize variables and build models over discrete domains. That is, the values for each node in the BBN should be categorized into the finite number of levels. How to discretize the variables is more difficult a question because the number and division points of the levels can make a notable difference in the complexity and precision of the resulting model. The bigger the number of levels, the more complex and precise the model is but more data are needed for it to construct probabilistic dependencies. In practice, 2–10 levels are reported in ecological studies (Uusitalo 2007). In this research, the concentrations of the seven hazardous substances are divided into five levels, as shown in Table 3.

Development of conditional probabilities

The difficulty of defining conditional probabilities arises when the relevant literature on probabilistic relationships between causes and effects is insufficient. In such situation, experts are usually able to use subjective judgment to assist in this task. In this paper, the determination of conditional probabilities suffers this difficulty due to the lack of sufficient information and therefore subjective judgment is exploited. In Taiwan, the morbidities of central nervous system diseases, dermatitis, kidney diseases, liver complaint, cardiovascular diseases, respiratory system diseases, and mortality for 20- to 59-year-old males are 11.36, 23.27, 1.11, 2.36, 3.57, 4.67, and 1.07 %, respectively (Taiwan Department of Health 2009), which are baseline situations. Subsequently, experts use their expertise to determine the morbidities and mortality for the worst situations.



Level	Toluene (ppm)	Trichloroethylene (ppm)	Trichloromethane (ppm)	Tetrachloroethylene (ppm)	Carbon monoxide (ppm)	Xylene (ppm)	Hydrogen sulfide (ppm)
Very low	0–200	0–100	0–70	0–100	0–70	0–200	0–20
Low	201-400	101-200	71-140	101-200	71-140	201-400	21-40
Medium	401-600	201-300	141-210	201-300	141–210	401-600	41-60
High	601-800	301-400	211-280	301-400	211-280	601-800	61-80
Very high	801-	401-	281-	401-	281-	801-	81–

Table 3 Five concentration levels of hazardous substances in sanitary sewerage

Therefore, morbidities of central nervous system diseases, dermatitis, kidney diseases, liver complaint, cardiovascular diseases, respiratory system diseases, and mortality for the worst situations are 70.8, 66.4, 27.8, 42.4, 35.7, 46.7, and 5.59 %, respectively. The rest conditional probabilities between baseline and worst situations are calculated based on an interpolative basis.

Software

A widely used software package, Hugin (HUGIN EXPERT A/S, Denmark), is employed to implement the BBN-HRA model. Its graphical user interfaces make it easy for users to build the network structure manually. It does offer the mechanism for learning CPTs from cases, however, the authors do not take advantage of it due to insufficient cases. Ultimately, the BBN-HRA model contains 14 nodes (variables), 22 arrows (causal relationships), 7 prior probabilities, 7 CPTs, and 950 conditional probabilities, as shown in Fig. 2. The BBN-HRA model is available free of charge from the corresponding author.

Results and discussion

Study area

Kaohsiung City is a coastal city with a population of about 150 million and becomes the transportation and commercial center in southern Taiwan. Besides, five industrial zones, two export processing zones, and several other important factories such as Taiwan Steel Corporation, Taiwan International Shipbuilding Corporation, Chinese Petroleum Corporation, etc., turn Kaohsiung City into an industrial center as well. However, accompanied by the rapid development of Kaohsiung City, environmental pollution seriously affects the life quality of the residents. Before 1989, the domestic sewage and industrial wastewater were directly or indirectly discharged into the river, threatening human and environmental health because the





Fig. 3 Kaohsiung City sanitary sewerage and nine sample points (Lin 2005)

sewage connection rate in this city was only 6.5 %. Therefore, at that time the Kaohsiung City Government launched the Kaohsiung City sewage system, as shown in Fig. 3. The sewage connection rate in this city was upgraded to 60.89 % in the end of 2010.

Lin (2005) analyzed the hazardous substances in the Kaohsiung City sanitary sewerage and took nine sampling points (from A to I, see Fig. 3) spreading over industrial, commercial, and residential districts. He considered two situations for each sample point, before ventilation and with ventilation for 15 min. To illustrate our approach, this

Table 4 Measurements of

	Stalidard (8-11 lifeall)	Belore ventilation			ventuation for 13 min		
		A	В	С	A	В	С
Toluene (ppm)	100	5.4	28.8	ND	0.0026	0.0138	ND
Trichloroethylene (ppm)	50	5.2	ND	ND	0.0025	ND	ND
Trichloromethane (ppm)	10	4.5	327.3	131.3	0.0022	0.1572	0.0630
Tetrachloroethylene (ppm)	50	21.0	223.9	103.3	0.0101	0.1755	0.0496
Carbon monoxide (ppm)	35	0.5	85.4	84.4	0.0002	0.0410	0.0405
Xylene (ppm)	100	10.7	12.3	7.5	0.0051	0.0059	0.0036
Hydrogen sulfide (ppm)	10	ND	4.1	ND	ND	0.0020	ND
	Toluene (ppm) Trichloroethylene (ppm) Trichloromethane (ppm) Tetrachloroethylene (ppm) Carbon monoxide (ppm) Xylene (ppm) Hydrogen sulfide (ppm)	Toluene (ppm)100Trichloroethylene (ppm)50Trichloromethane (ppm)10Tetrachloroethylene (ppm)50Carbon monoxide (ppm)35Xylene (ppm)100Hydrogen sulfide (ppm)10	Toluene (ppm)1005.4Trichloroethylene (ppm)505.2Trichloroethylene (ppm)104.5Tetrachloroethylene (ppm)5021.0Carbon monoxide (ppm)350.5Xylene (ppm)10010.7Hydrogen sulfide (ppm)10ND	Standard (3-n mean)Detote ventural A B Toluene (ppm)1005.428.8Trichloroethylene (ppm)505.2NDTrichloroethylene (ppm)104.5327.3Tetrachloroethylene (ppm)5021.0223.9Carbon monoxide (ppm)350.585.4Xylene (ppm)10010.712.3Hydrogen sulfide (ppm)10ND4.1	Standard (8-n mean) Detote Ventuation A B C Toluene (ppm) 100 5.4 28.8 ND Trichloroethylene (ppm) 50 5.2 ND ND Trichloroethylene (ppm) 10 4.5 327.3 131.3 Tetrachloroethylene (ppm) 50 21.0 223.9 103.3 Carbon monoxide (ppm) 35 0.5 85.4 84.4 Xylene (ppm) 100 10.7 12.3 7.5 Hydrogen sulfide (ppm) 10 ND 4.1 ND	Jandard (o-n mean)Defore ventuationventuation A B C A Toluene (ppm)100 5.4 28.8 ND 0.0026 Trichloroethylene (ppm)50 5.2 NDND 0.0025 Trichloroethylene (ppm)10 4.5 327.3 131.3 0.0022 Tetrachloroethylene (ppm)50 21.0 223.9 103.3 0.0101 Carbon monoxide (ppm)35 0.5 85.4 84.4 0.0002 Xylene (ppm)100 10.7 12.3 7.5 0.0051 Hydrogen sulfide (ppm)10ND 4.1 NDND	ABCABToluene (ppm)100 5.4 28.8 ND 0.0026 0.0138 Trichloroethylene (ppm)50 5.2 NDND 0.0025 NDTrichloroethylene (ppm)10 4.5 327.3 131.3 0.0022 0.1572 Tetrachloroethylene (ppm)50 21.0 223.9 103.3 0.0101 0.1755 Carbon monoxide (ppm)35 0.5 85.4 84.4 0.0002 0.0410 Xylene (ppm)100 10.7 12.3 7.5 0.0051 0.0029 Hydrogen sulfide (ppm)10ND 4.1 NDND 0.0020

paper only adopts three sampling points (A, B, and C) where the concentrations of the hazardous substances are listed in Table 4. Sampling point A lies in an interception station, in which the concentration of tetrachloroethylene is relatively high but still lower than the standard value. Sampling points B and C are situated in a residential district but close to an export processing zone; therefore, the concentrations in trichloromethane and tetrachloroethylene are extraordinary high. After 15-min ventilation, the concentrations of hazardous substances in the three sample points drop dramatically, as shown in Table 4.

HRA through BBN

The information of the three sample points is inputted into the BBN-HRA model. For example, the results of sampling point B before ventilation are estimated as follows: central nervous system diseases of 41.10 %, dermatitis of 23.27 %, kidney diseases of 14.47 %, liver complaint of 22.39 %, cardiovascular diseases of 5.58 %, respiratory system diseases of 4.67 %, and mortality of 1.90 %, as demonstrated in Fig. 4. After 15-min ventilation, the morbidity and mortality rates in this sample point will decrease to the baseline conditions; that is, central nervous







Table 5 HRA analysis of case study

	Standard (8-h mean)	Before ventilation			Ventilation for 15 min		
		A	В	С	A	В	С
BBN-HRA							
Central nervous system diseases (%)	11.36	11.36	41.10	18.80	11.36	11.36	11.36
Dermatitis (%)	23.27	23.27	23.27	23.27	23.27	23.27	23.27
Kidney diseases (%)	1.11	1.11	14.47	5.56	1.11	1.11	1.11
Liver complaint (%)	2.36	2.36	22.39	9.04	2.36	2.36	2.36
Cardiovascular diseases (%)	3.57	3.57	5.58	5.58	3.57	3.57	3.57
Respiratory system diseases (%)	4.67	4.67	4.67	4.67	4.67	4.67	4.67
Mortality (%)	1.07	1.07	1.90	1.42	1.07	1.07	1.07
HQ							
Toluene	1.00	0.05	0.29	0.00	0.00	0.00	0.00
Trichloroethylene	1.00	0.10	0.00	0.00	0.00	0.00	0.00
Trichloromethane	1.00	0.45	32.73	13.13	0.00	0.02	0.01
Tetrachloroethylene	1.00	0.42	4.48	2.07	0.00	0.00	0.00
Carbon monoxide	1.00	0.01	2.44	2.41	0.00	0.00	0.00
Xylene	1.00	0.01	0.01	0.01	0.00	0.00	0.00
Hydrogen sulfide	1.00	0.00	0.41	0.00	0.00	0.00	0.00
Central nervous system diseases	1.00	0.50	33.02	13.13	0.00	0.02	0.01
TOSHI							
Dermatitis	1.00	0.16	0.29	0.00	0.00	0.00	0.00
Kidney diseases	1.00	0.92	37.50	15.20	0.00	0.02	0.01
Liver complaint	1.00	0.97	37.21	15.20	0.00	0.02	0.01
Cardiovascular diseases	1.00	0.18	2.74	2.42	0.00	0.00	0.00
Respiratory system diseases	1.00	0.01	0.42	0.01	0.00	0.00	0.00

system diseases of 11.36 %, dermatitis of 23.27 %, kidney diseases of 1.11 %, liver complaint of 2.36 %, cardiovascular diseases of 3.57 %, respiratory system diseases of 4.67 %, and mortality of 1.07 %. The entire outcomes of the morbidity and mortality rates for the three sample points are listed from the 10th to 16th rows of Table 5. In these rows, the results higher than the standard values are denoted by single-underlines; obviously, they occur only in sample points *B* and *C* before ventilation, including central nervous system diseases, kidney diseases, liver complaint, and cardiovascular diseases (Figs. 4 and 5).

Discussion: comparison of BBN-HRA with TOSHI

In risk characterization of noncarcinogenic substances, the HRA can be evaluated by the HQ, the ratio of the intake of a hazardous substance to its reference dose. Multiple hazardous substances may affect the same organ (or organ system) causing joint effect, and hence, the TOSHI sums the HQ scores of multiple substances that have joint effect on a specific organ (US EPA 2001). In this case study, Fig. 3 shows that the target organ systems of toluene are cardiovascular and respiratory systems;



trichloroethylene: dermatitis, liver complaint, and cardiovascular system; trichloromethane: central nervous system, kidney, and liver; tetrachloroethylene: kidney and liver; carbon monoxide: cardiovascular system; xylene; cardiovascular system and respiratory system; hydrogen sulfide: respiratory system. On the basis of the toluene, trichloroethylene, trichloromethane, tetrachloroethylene, carbon monoxide, xylene, and hydrogen sulfide concentrations described in Sect. 3.1, their HQ and TOSHI scores can be computed, as shown in the 17th-29th rows of Table 5. The HQ and TOSHI scores reveal that trichloromethane, tetrachloroethylene, and carbon monoxide in sample points B and C before ventilation probably cause central nervous system diseases, kidney diseases, liver complaint, and cardiovascular diseases, and adverse noncancer health effects; the rest hazardous substances will not induce any adverse noncancer effect. The comparisons between the TOSHI and the BBN-HRA are discussed in the following.

The HQ score for non-critical organs is overestimated because it is derived from the reference dose of the most critical organ with the lowest NOAEL. The same HQ score is assigned to all (critical or non-critical) organs due to the





general absence of organ-specific reference doses. In Table 5, the HQ score of trichoromethane in sample point B before ventilation is transferred to central nervous system, kidney, and liver simultaneously although it is known that the critical effects occur in central nervous system. On the other hand, the BBN-HRA is developed according to epidemiological studies, the actual hospital attendance records, and expert experiences to elaborate the respective morbidity and mortality rates of all the related diseases under certain pollution concentrations. The results of BBN-HRA in sample point B before ventilation in Table 5 reveal that the critical effect of trichoromethane is central nervous system diseases while its non-critical effects are kidney and liver diseases.

The above additive assumption in TOSHI indicates that it is unlikely for a TOSHI score below 1.0 to result in adverse noncancer health effects over a lifetime of exposure. whereas a score above 1.0 does not necessarily suggest a likelihood of adverse effects (US EPA 2001). In Table 5, the TOSHI score of cardiovascular system in sample point *B* before ventilation is 2.74, which makes it difficult to decide whether it will cause illness or not. However, through BBN-HRA calculation an increase of cardiovascular disease from 3.57 to 5.58 % is more informative.

Conclusion

This study proposed a BBN-HRA model for sewer workers with features including the representation of probabilistic relationships between hazardous substances and adverse human effects through the graph structures of the BBN, the construction of dose-response relationships by CPTs of the BBN, and the capability of predicting morbidity and mortality rates of the related diseases with the inference mechanism of the BBN. This BBN-HRA model can address some problems of applying the TOSHI to HRA. Its graph structure can pinpoint the relationships between hazardous substances and the induced diseases. Moreover, on the basis of epidemiological studies, actual hospital attendance records and expert experiences, the BBN-HRA can specifically identify the probability of each induced disease under certain concentration of a hazardous substance, which can overcome the problem of overestimation of HQ for non-critical organs. The BBN-HRA model provides more concrete information on the morbidity and mortality rates of all the related diseases, thus reducing the uncertainty in the TOSHI.

The BBN-HRA model was demonstrated by a practical case study, which shows that with no ventilation in sewers, the probabilities for the workers in the worst situation



(sample point *B*) to contract central nervous system diseases, dermatitis, kidney diseases, liver complaint, cardiovascular diseases, respiratory system diseases, and mortality were 41.10, 23.27, 14.47, 22.39, 5.58, 4.67, and 1.90 %, respectively. However, a 15-min ventilation reduced the abovementioned probabilities to the baseline condition, that is, 11.36, 23.27, 1.11, 2.36, 3.57, 4.67, and 1.07 %, respectively.

The authors suffered several difficulties in applying the BBN to HRA and they still need further endeavor to solve. The first one was to discretize appropriately the variables because the bigger the number of discretization, the more complex and precise the CPTs are, but more data are needed. In this paper, five concentration levels of hazardous substances were adopted but they required more solid study in future research. The second difficulty was to gather sufficient epidemiological studies to avoid subjective judgments. The third difficulty came from insufficient studies on the joint effects of multiple pollutants, which compelled us to use expert experiences in aggregating conditional probabilities. Indeed, the last two difficulties are not due to the model itself, but for the lack of relevant epidemiological studies to support this model. If these difficulties can be overcome, the BBN will be very beneficial in HRA.

Acknowledgments The authors would like to thank the National Science Council of the Republic of China (Taiwan) for financially supporting this research under Contract NSC 99-2221-E-131-010-MY2.

References

- Batayneh AT (2012) Toxic (aluminum, beryllium, boron, chromium and zinc) in groundwater: health risk assessment. Int J Environ Sci Technol 9(1):153–162
- California EPA (California Environmental Protection Agency) (2001) A Guide to Health Risk Assessment. Available at http://oehha. ca.gov/pdf/HRSguide2001.pdf

- Dawsey WJ, Minsker BS, Van Blaricum VL (2006) Bayesian belief networks to integrate monitoring evidence of water distribution system contamination. Water Res Plan Manag 132:234–241
- Lai CH, Lin CH, Yeh SH, Huang LJ (2004) Air pollutant exposure and risk assessment for sewer workers. Environ Inform Archives 2:405–412
- Liao Y, Wang J, Guo Y, Zheng X (2010) Risk assessment of human neural tube defects using a Bayesian belief network. Stoch Env Res Risk Assess 24:93–100
- Lin GX (2005) Risk assessment for sewer workers in Kaohsiung city, master's thesis, Fooyin University, Kaohsiung
- Liu KFR, Lu CF, Chen CW, Shen YS (2012) Applying Bayesian belief networks to health risk assessment. Stoch Env Res Risk Assess 26:451–465
- Newton AC (2010) Use of a Bayesian network for Red Listing under uncertainty. Environ Model Softw 25:15–23
- Pearl J (1988) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann, California
- Shyam R, Kalwania GS (2012) Health risk assessment of fluoride with other parameters in ground water of Sikar city (India). Environ Earth Sci 65:1275–1282
- Tahir MW, Mumtaz MW, Tauseef S, Sajjad M, Nazeer A (2012) Monitoring of cotton dust and health risk assessment in smallscale weaving industry. Environ Monit Assess. doi:10.1007/ s10661-011-2309-y
- Taiwan Department of Health (2009) The annual rate of outpatient visits in 2009. Available at http://www.doh.gov.tw
- Ticehurst JL, Curtis A, Merritt WS (2010) Using Bayesian Networks to complement conventional analyses to explore landholder management of native vegetation. Environ Model Softw 26:52– 65
- US EPA (US Environmental Protection Agency) (2001) National-Scale Air Toxics Assessment for 1996: Draft for EPA Science Advisory Board Review. Available at http://www.epa.gov/ ttn/atw/sab/sabrev.html#A1
- Uusitalo L (2007) Advantages and challenges of Bayesian networks in environmental modeling. Ecol Model 203:312–318
- Xue ZJ, Liu SQ, Liu YL, Yan YL (2012) Health risk assessment of heavy metals for edible parts of vegetables grown in sewageirrigated soils in suburbs of Baoding City, China. Environ Monitor Assess. doi:10.1007/s10661-011-2204-6
- Yeh SH, Lai CH, Lin CH et al (2011) Estimating cancer risk Increment from air pollutant exposure for sewer workers working in an industrial city. Aerosol Air Quality Res 11(2): 120–127

