Popul Environ (2014) 36:219–233
DOI 10.1007/s11111-014-0212-5

RESEARCH BRIEF

An empirical model to predict arsenic pollution affected life expectancy

S. R. Samadder · D. Nagesh Kumar · N. M. Holden

Published online: 2 April 2014

Abstract A robust, globally implementable and simple empirical model to predict the arsenic pollution affected life expectancy using a stepwise regression was developed. Life expectancy calculated using a life table technique requires crude death rates data that are not available for small administrative units, complex calculations and does not consider socioeconomic parameters. Hence, a model was needed to forecast the impact of arsenic pollution and socioeconomic parameters on life expectancy for locations with limited data availability. A linear multiple regression technique was used to develop an empirical model to predict arsenic pollution affected life expectancy at birth. The model was calibrated using nine arsenic polluted administrative blocks of district Murshidabad, West Bengal, India and tested independently for three other arsenic polluted blocks of the same district. The R^2 values for the plot of actual versus predicted life expectancy at birth were 0.98 for calibration, testing and independent validation. The model is complementary to the life table technique and offers a means to assist planning by public health engineers and health policy makers to mitigate arsenic pollution on a community priority basis.

Keywords Arsenic pollution · Empirical model · Groundwater · Life expectancy · Model validation · Potable water · Regression analysis

S. R. Samadder (✉)
Department of Environmental Science and Engineering, Indian School of Mines, Dhanbad 826004, India
e-mail: sukh_samadder@yahoo.co.in

D. Nagesh Kumar
Department of Civil Engineering, Indian Institute of Science, Bangalore 560012, India

N. M. Holden
Agriculture and Food Science Centre, UCD School of Biosystems Engineering, University College Dublin, Belfield, Dublin 4, Ireland

Introduction

The reported incidence of chronic arsenic poisoning in Bangladesh and West Bengal (in India) from long term use of groundwater with high arsenic concentration has come as a shock to drinking water suppliers, users, and scientists (Welch and Stollenwerk 2003). As many as 35 million people have been exposed to groundwater with an arsenic concentration >0.05 mg/L (maximum permissible level in India and Bangladesh) and about 57 million people are exposed to concentrations >0.01 mg/L (WHO maximum permissible level) in Bangladesh (Gaus et al. 2001). High arsenic concentrations have been observed in groundwater in an area covering more than 23,000 km^2 of the lower Ganga delta stretching from India to Bangladesh (Central Ground Water Board 1999).

Life expectancy as an indicator of population health (Lai et al. 2000) computed by life table techniques summarizes the death process in a given population (Das 2000). In practice, the mortality data of a current population are used to construct a hypothetical cohort, experiencing mortality estimated from that of the observed population (Elandt-Johnson and Johnson 1980). In recent decades, life expectancy has increased more rapidly in developing countries than developed countries (Lai et al. 2000). Living longer may not necessarily imply a better quality of life (Rogers et al. 1990), thus quality of life can be used to adjust expected life measures (Mutafova et al. 1997; Rogers et al. 1989; Valkonen et al. 1997). Analyses of adjusted life expectancy for developed countries have been undertaken (Nusselder et al. 1996; Crimmins et al. 1997), but detailed analysis has not been reported due to lack of adequate data.

Carcinogenicity of arsenic depends on (a) food practice (b) concentration (c) period of exposure (d) nutritional status (e) genetic variation of the population in response to exposure (f) immunity level of the population (g) age and (h) sex of the exposed individuals (World Health Organization 1984). Samadder (2010) forecast the impact of groundwater arsenic on life expectancy at birth using these factors and life table methods. Smith et al. (2000) reported that the ingestion of arsenic, both from water supplies and pharmaceutical preparations causes skin cancer and internal cancers. National Research Council (2004) calculated that males whose daily consumption of water containing 0.05 mg/L of arsenic have about a one in 1,000 risk of developing bladder cancer, and Brown et al. (1989) found lifetime risk of developing skin cancer is 1.3 in 1,000 for males and 0.6 in 1,000 for females per microgram of arsenic consumption per day. Smith et al. (1992, 2000) established that at 0.05 mg/L of arsenic, the lifetime risk of dying from liver, lungs, kidney, or bladder cancer from drinking 1 L/day can be as high as 13.4 per 1,000 persons and at 0.50 mg/L, it can be as high as 134 per 1,000 persons. Tseng et al. (1968), Chen et al. (1988), Wu et al. (1989), Abernathy et al. (1997) and Samadder (2010) all reported a linear dose–response relationship between the arsenic concentration in drinking water and mortality rate.

Life expectancy calculated using life table techniques requires crude death rate data that are not available for small administrative units, complex calculations and does not consider socioeconomic parameters (Samadder 2010). It was therefore proposed that a comprehensive model of life expectancy at birth that considers arsenic pollution and socioeconomic parameters, with a limited data requirement and the potential to help understand and explain the results was needed. The objective of this paper was to design, parameterize, calibrate, and test an empirical model of arsenic affected life expectancy for an area with serious problems and limited data availability. The study area was the district of Murshidabad, located in the Gangetic plains of West Bengal, India. It is severely polluted by arsenic in groundwater (Samadder 2005, 2010; Samadder and Subbarao 2007), and there is little water treatment and few distribution networks. The population is poorly educated and in general does not understand the consequences of directly consuming contaminated well water obtained using easily installed, low-cost hand pumps.

Materials and methods

Murshidabad district is one of the worst affected areas in the world subject to arsenic pollution. The district has 26 administrative blocks, and a previous analysis showed that 22 blocks had arsenic concentrations above 0.05 mg/L in groundwater (Samadder 2005). The 12 peripheral blocks (shaded gray in Fig. 1) were chosen as the study area for this work based on the severity of arsenic pollution in groundwater, proximity to river Ganges (called Padma in Bangladesh), and reported arsenicosis cases to the area hospitals. The 12 arsenic affected administrative blocks (study area) are subdivided into 590 mouzas (smallest administrative unit in the region, Fig. 2) occupies 0.201 million ha and support 2.3 million people. For each mouza, there are several tube wells for which the arsenic concentration values were known. The number of wells in a mouza depends on the population density and total area of the mouza. The tube wells are evenly distributed in populated areas. All the tube wells are used for drinking water and other domestic purposes. Most of the tube wells are community based, but people do not necessarily take water from a single tube well as they move from one place to other (within a mouza in most cases) to earn their daily wages. The population was assumed to be exposed to the average concentration of arsenic within an area (rather than median concentration as used by Wong et al. (1997). Other data sets were collected at the mouza-scale, and the number of mouza represents the number of samples in the present study. Out of the 12 blocks, nine administrative blocks (Suti-1, Raghunathganj-2, Lalgola, Bhagabangola-1, Bhagabangola-2, Raninagar-1, Raninagar-2, Jalangi, and Beldanga-1) (Fig. 1) consisting of 402 mouzas ($n = 402$) were selected to calibrate the model and the remaining three blocks (Domkal, Hariharpara, and Nawda) (Fig. 1) consisting of 188 mouzas ($n = 188$) were used as independent test data. The 12 arsenic affected administrative blocks are located between 23°43′30″N–24°50′20″N and 87°49′17″E to 88°46′00″E.

The age-specific crude death rate data at State level and distribution of population data at mouza-scale by age group for the year 1999 were collected from the Registrar General of India, New Delhi, and Census of India, Directorate of Census Operations, West Bengal, respectively. Maps showing mouza boundaries and other topographical information were acquired from the District Land and Land Reforms

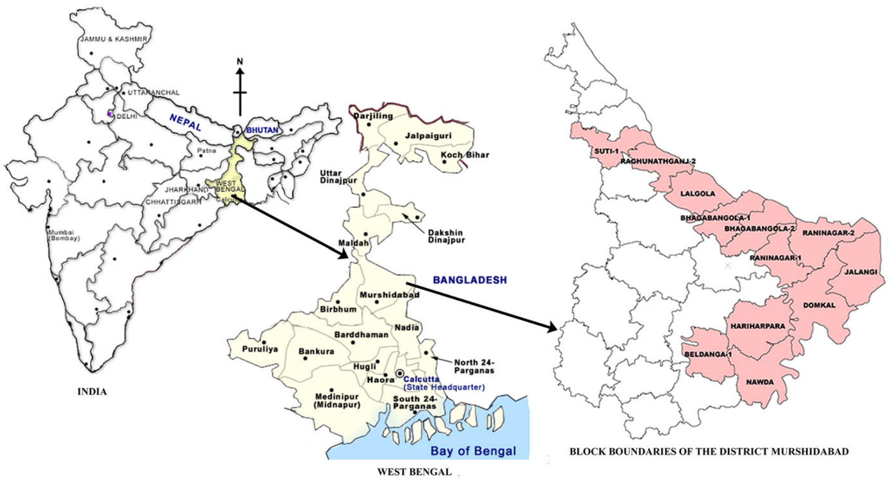

Fig. 1 Location of the district Murshidabad and its administrative blocks in West Bengal, India

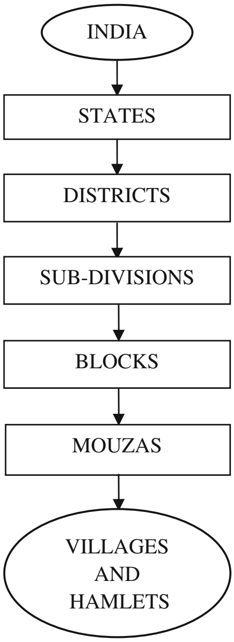

Fig. 2 Administrative hierarchy in the study region

Office for Murshidabad. The arsenic distribution in groundwater for the study area was derived from 28,357 groundwater samples collected and analyzed by the Public Health Engineering Department (PHED) for Murshidabad, West Bengal.

The mouza scale arsenic pollution affected life expectancy at birth was calculated using Newell's method (Samadder 2010; Newell 1988) and was considered as the dependent or predicted variable in the model (Y). As the study area is rural and dependent on agricultural economy, people living there have moderate facilities and arsenic pollution is of major concern, yet data availability is limited. Those data that could be collected (from the local Government Offices, such as Office of the District Magistrate and Public Health Engineering Department of the study area) and that might influence the life expectancy were collected at mouza scale as independent or predictor variables. These were *area* in ha (X_1), *child population* (X_2), *average arsenic concentration in groundwater* in mg/L (X_3), *average depth to average arsenic concentration* in meters (X_4), *total population using the water* (X_5), *literacy rate* (X_6), *risk due to consumption of local arsenic contaminated water* as a percentage (X_7) (calculated based on the findings of Smith et al. 1992, 2000); and evidence of a linear dose–response relationships, arsenic concentration values between 0.05 and 0.50 mg/L were divided into ranges in multiples of 0.05 and the corresponding risk factors in multiples of 1.34 (percent risk for arsenic concentration from 0 to 0.05 mg/L was considered as 0 as this is the drinking water standard for India, percent risk for arsenic concentration range 0.05–0.10 was considered as 2.68 % (2×1.34), percent risk for arsenic concentration range 0.10–0.15 was considered as 4.02 (3×1.34) and so on), and *forecast death rate due to consumption of arsenic contaminated water* (X_8) (calculated based on percent risk and total number of people using such water). A dataset of 402 samples (representing the number of mouzas in nine administrative blocks) (Fig. 1) of Y and $X_1 \ldots X_8$ was compiled.

The model was developed in SPSS version 10 (Landau and Everitt 2004) using stepwise regression to select the possible subset of X_i to predict the dependent variable (Y). Stepwise regression was chosen because it avoids irrational coefficients as statistical criteria used to select the predictor variables have high intercorrelation (McCuen and Snyder 1986). Forward stepwise regression with deletion was used in a three step process: (1) partial F tests for insertion, (2) total F tests for model significance and (3) partial F tests for deletion. Further details on partial and total F tests can be found in McCuen and Snyder (1986).

Step (1): Partial F test for insertion

The partial F values for all predictor variables not included in the equation were computed, and the variable with the largest partial F value was selected to enter the equation. In the first step, the partial correlations were compared with the predictor-predicted correlations, and the predictor variable that had the largest predictor-predicted correlation was added first. If $F < F_\alpha$, the variable was not statistically significant at $\alpha = 0.05$ and predictor variable was not significantly related to the predicted variable. If $F > F_\alpha$, the variable was statistically significant, and the variable was included in the model, which was assessed by the total F test (step 2).

Step (2): Total F test

An equation that included all predictor variables that were inserted (and not subsequently deleted) was obtained, and a total F value calculated and compared with the critical F_α. If $F < F_\alpha$, the equation attained from the previous iteration was used as the final regression model, but if $F > F_\alpha$, the entire model was significant, and all the predictor variables that were in the equation were tested for significance; control then proceeded to step 3.

Step (3): Partial F test for deletion

Partial F values for all predictor variables that were included in the equation were computed. The variable that had the smallest F value was compared with the critical F value, F_α. If $F < F_\alpha$, the predictor variable being tested was insignificant and deleted from the equation; control then passed to the total F test (step 2). Whereas, if $F > F_\alpha$, all predictor variables included in the equation were statistically significant; control passed to the partial F tests for insertion (step 1).

The F value to include a predictor variable in the model was 3.84 ($\alpha = 0.05$) and that to remove a predictor variable was 2.71 ($\alpha = 0.1$). The model was independently tested for the blocks excluded from the calibration of the model [Domkal ($n = 87$), Hariharpara ($n = 62$), and Nawda ($n = 39$)]. The independent variables (X_i) chosen for the model development were tested for a linear relationship with the dependent variable (Y) using Pearson's correlation coefficient following the Kolmogorov–Smirnov test for normality. The uncertainty of the model was quantified by the coefficient of determination (R^2), standard error of estimate, and residual analysis. The model was cross-validated using the predicted arsenic pollution affected life expectancy at birth from the model and arsenic pollution affected life expectancy at birth calculated using life table technique (Samadder 2010; Newell 1988) for the nine administrative blocks used for model development. The dataset (both life expectancy values from the model and corresponding life expectancy values using life table technique) of the nine administrative blocks was randomly ordered and plotted to test the model taking all samples of the nine blocks, the first 80 % of data points, and the remaining 20 % of data points separately.

Results

Four of the eight independent variables were found to generate the best model (Eq. 1). In order of contribution, the predictor variables were: X_7 (risk due to consumption of arsenic contaminated water); X_3 (the average arsenic concentration in groundwater); X_8 (forecast death rate due to consumption of arsenic contaminated water); and X_5 (total population using the water). The best prediction model was as follows:

$$Y = 65.143 - 3.457X_7 + 21.909X_3 + 0.002933X_8 - 0.00003139X_5 \qquad (1)$$

The summary statistics for the model are presented in Tables 1, 2 and 3. The assumptions of normality and linearity were both met. The maximum absolute difference between the values of the cumulative distribution of random sample and the cumulative function of normal probability distribution was 0.447 and the corresponding Kolmogorov–Smirnov Z value specified was 10.850, so the data were normally distributed [also seen in the normal probability plot (Fig. 3)]. All the independent variables were tested for their correlations with each other (test of independence) and with the dependent variable (Tables 4, 5), and it was observed that out of the four variables used in the model (Eq. 1), X_3 had significant correlation with X_7 and X_8. Similarly, X_5 had significant correlation with X_1 and X_2. Of the four independent variables, three were linearly related to the dependent variable but X_5 was not.

When the independent variables were used to model arsenic pollution affected life expectancy at birth using all data (Fig. 4a), an excellent fit was achieved ($n = 402$, $R^2 = 0.9840$). Cross calibration indicated no deterioration using a random 80 % of data ($n = 322$, $R^2 = 0.9841$) for calibration (Fig. 4b) and the remaining 20 % for verification ($n = 80$, $R^2 = 0.9839$) (Fig. 4c). When the model was independently tested with data from three administrative blocks not included in model calibration, an excellent fit was achieved in each case: Domkal ($n = 87$, $R^2 = 0.9829$) (Fig. 5a), Hariharpara ($n = 62$, $R^2 = 0.9863$) (Fig. 5b), and Nawda ($n = 39$, $R^2 = 0.9896$) (Fig. 5c).

Discussion

Four of the eight independent variables were found insignificant to generate the model. The independent variables excluded from the model were as follows: area (X_1), child population (X_2), average depth to average arsenic concentration (X_4), and literacy rate (X_6). In order of contribution, the necessary predictor variables were as follows: X_7 (risk due to consumption of arsenic contaminated water); X_3 (average arsenic concentration in groundwater); X_8 (forecast death rate due to consumption of arsenic contaminated water); and X_5 (total population using the water). Though the coefficient for X_5 (Eq. 1) was very small, the value of X_5 was as high as 28,028 (population in a mouza).

The arsenic pollution affected life expectancy values that were calculated using arsenic pollution affected crude death rates data derived after forecasting the number of death cases due to consumption of arsenic polluted water. The average arsenic in groundwater (X_3) in the final model reflected the amount of arsenic ingested by the population. If the value of X_3 increases then the quantity of arsenic ingested and impact on life expectancy will increase due to the linearly related impact of carcinogenicity of arsenic in the human body. Total population using the water (X_5) governed the number of people exposed to arsenic pollution through drinking water. In this study, X_5 represented how many people would risk death and the arsenic pollution affected life expectancy must depend on how many people die due to arsenic consumption. Similarly, X_7 (percent risk of death) calculated how many people would die out of the total exposed population. The forecast death rate due to consumption of arsenic contaminated water (X_8) modified the crude death

Table 1 Stepwise development of the multiple linear regression model to the point where all significant independent variables were included

Step	R	R^2	Adjusted R^2	Std. Error of estimate	Change statistics			
					F change	R^2	v_1	v_2
Performance with reference to (e)								
1	0.988 (a)	0.975	0.975	0.6801	23,244.332	0.975	1	400
2	0.992 (b)	0.983	0.983	0.5627	271.974	0.008	1	399
3	0.992 (c)	0.984	0.984	0.5514	25.384	0.001	1	398
4	0.992 (d)	0.984	0.984	0.5440	16.845	0.000	1	397

(a) Predictors: X_7

(b) Predictors: X_7, X_3

(c) Predictors: X_7, X_3, X_8

(d) Predictors: X_5, X_3 X_8, X_5

(e) Dependent variable: Y

Table 2 Analysis of variance (ANOVA) or the stepwise development of the multiple linear regression model

Step	Sum of squares	*df*	Mean square	F	Significance (95 %)
ANOVA (*e*)					
1					
Regression	10,751.208	1	10,751.208	23,244.332	0.000 (a)
Residual	271.968	400	0.463		
Total	11,023.176	401			
2					
Regression	10,837.321	2	5,418.660	17,114.112	0.000 (b)
Residual	185.856	399	0.317		
Total	11,023.176	401			
3					
Regression	10,845.037	3	3,615.012	11,891.818	0.000 (c)
Residual	178.139	398	0.304		
Total	11,023.176	401			
4					
Regression	10,850.023	4	2,712.506	9,164.231	0.000 (d)
Residual	173.153	397	0.296		
Total	11,023.176	401			

df degrees of freedom

(a) Predictors: X_7

(b) Predictors: X_7, X_3

(c) Predictors: X_7, X_3, X_8

(d) Predictors: X_7, X_3, X_8, X_5

(e) Dependent variable: Y

Table 3 Values of intercept and slope coefficients at different steps of stepwise regression

Step	Unstandardized coefficients		Standardized coefficients	t	95 % Confidence interval for B		Correlations	
	B	Std. Error	Beta		Lower bound	Upper bound	Partial	Part
Analysis with reference to (*e*)								
1								
(Constant)	65.515	0.032		2,020.662	65.451	65.578		
X_7	−2.814	0.018	−0.988	−152.461	−2.850	−2.778	−0.988	–
2								
(Constant)	64.972	0.042		1,530.128	64.888	65.055		
X_7	−3.316	0.034	−1.164	−97.374	−3.383	−3.249	−0.970	–
X_3	23.428	1.421	0.197	16.492	20.638	26.218	0.563	0.088
3								
(Constant)	64.988	0.042		1,557.533	64.906	65.069		
X_7	−3.366	0.035	−1.181	−96.717	−3.434	−3.297	−0.970	–
X_3	22.950	1.395	0.193	16.449	20.210	25.690	0.562	0.086
X_8	1.228E−	0.000	0.034	5.038	0.001	0.002	0.204	0.026
4								
(Constant)	65.073	0.046		1,409.709	64.983	65.164		
X_7	−3.402	0.035	−1.194	−95.936	−3.472	−3.332	−0.970	–
X_3	23.459	1.382	0.197	16.971	20.744	26.174	0.574	0.088
X_8	1.848E−	0.000	0.051	6.507	0.001	0.002	0.260	0.034
X_5	–	0.000	−0.025	−4.104	0.000	0.000	−0.167	–

(e) Dependent variable: Y

rate for arsenic pollution. All the four included variables in the model are mechanistically significant so were retained for prediction purposes.

The excluded variable area (X_1) should not be mechanistically linked with life expectancy (Y), and child population (X_2) was strongly correlated with total population (X_5) so only X_5 was needed in the model. Average depth to average arsenic concentration (X_4) could be a determining factor for the concentration of arsenic in groundwater in Bangladesh and West Bengal, but this was not the case in the study area. The literacy rate (X_6) might have explained population behavior if areas with higher literacy rates consumed less contaminated water, but no relationship was found for the study area.

Age-specific crude death rates data are generated from age-specific death rates observed in representative samples collected from different areas. The analysis did not consider specific cause of death observed in the samples used to generate the age-specific crude death rates using life table techniques. Some developing nations are improving life expectancy by improving medical facilities and food security, but impact of environmental pollution on life expectancy is not a separate focus during surveys. Arsenic pollution in groundwater is a serious issue of concern in the study area and in many developing countries including Bangladesh. The model developed

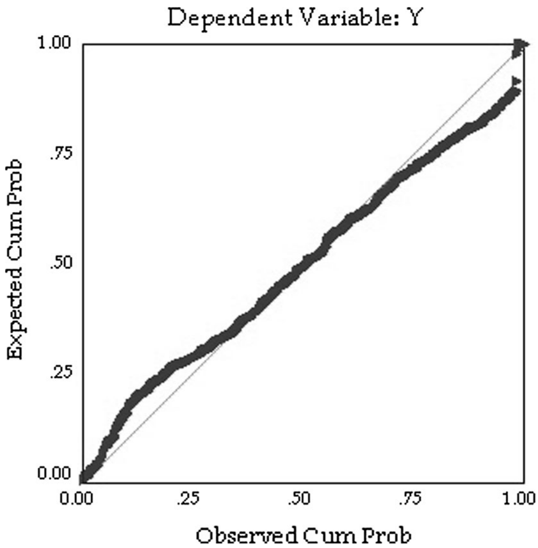

Fig. 3 Normal probability plot

will be used to help understand the impact of arsenic pollution on life expectancy for small communities varying from those using only one tube well for their drinking water supply through perhaps to districts. The calculated life expectancy for the affected areas using the model will provide information about the health status of the people concerned to facilitate alternative water supply schemes (free from arsenic pollution), allocation of arsenic treatment units or health care facilities on a priority basis depending on severity of the problem (i.e., forecast reduction in life expectancy). For example: the reduction in life expectancy for the block Lalgola (one of the twelve blocks of the study area) varied from 0 to 18.43 years. The resources should be allocated first to the area with a 18.43 years of life expectancy reduction.

The model is also useful for understanding the impact of cause-specific deaths (arsenic ingestion for the present study) on life expectancy for a small community. Life expectancy calculated at all levels of the administrative hierarchy provides a means for allocating resources at national and local government levels to improve the life expectancy, and the model will facilitate resource allocation for delineated arsenic polluted areas (particularly in developing nations such as India and Bangladesh), which is a difficult task when resources are limited.

Using the model to mitigate arsenic pollution for affected communities on a priority basis will help focus public health engineering and health policy where it is needed. For example: Jyotkanai and Jhauberia are two mouza of the block Domkal where average arsenic concentrations were 0.074 and 0.211 mg/L, respectively; and the corresponding total population exposed to arsenic pollution were 5,146 and 1,219, respectively. The life expectancy calculated using the model for mouza Jyotkanai was 57.74 years (reduced by 7.93 years) and that of mouza Jhauberia was

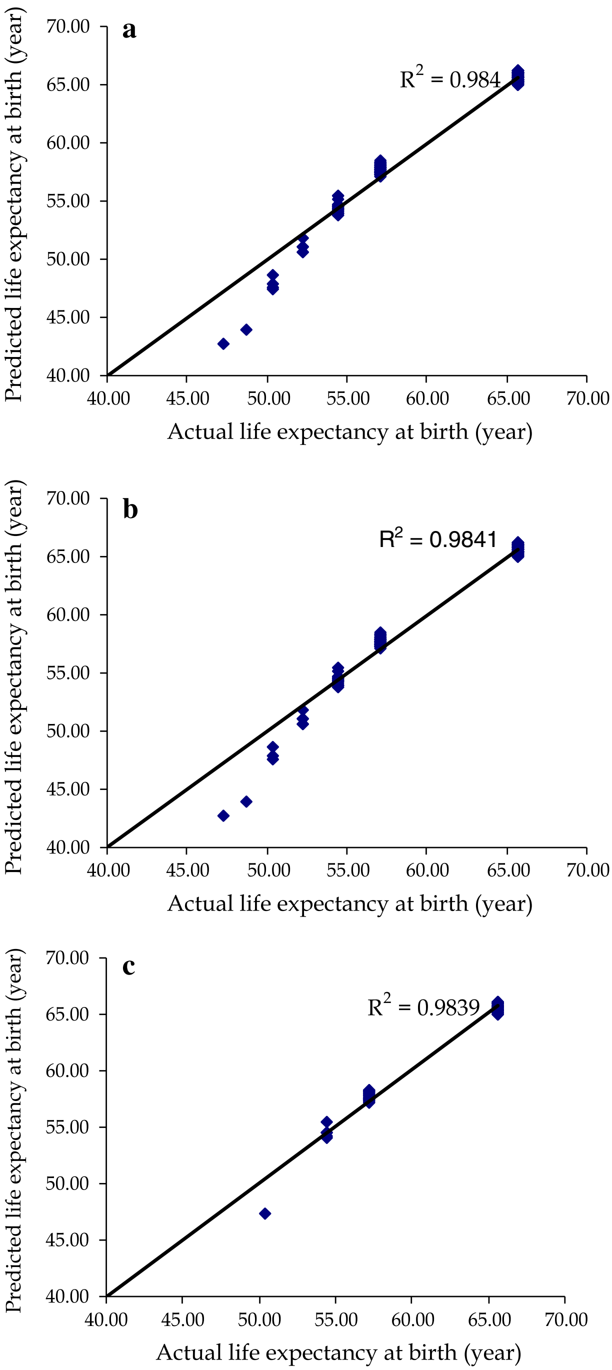

Fig. 4 **a** Relation between actual and predicted life expectancy at birth considering all samples. **b** Relation between actual and observed life expectancy at birth taking first 80 % of the samples randomly. **c** Relation between actual and observed life expectancy at birth taking remaining 20 % of the sample randomly

Table 4 Pearson's correlation analysis for linear dependency among independent variables ($n = 402$)

Independent variables	Independent variables							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1								
Pearson correlation	1.000	0.608	0.000	0.038	0.653	0.029	−0.015	0.172
Significance (2-tailed)	–	0.000	0.997	0.355	0.000	0.476	0.724	0.000
X_2								
Pearson correlation	0.608	1.000	0.103	0.084	0.985	0.059	0.064	0.449
Significance (2-tailed)	0.000	–	0.013	0.041	0.000	0.151	0.119	0.000
X_3								
Pearson correlation	0.000	0.103	1.000	−0.030	0.110	0.075	0.894	0.583
Significance (2-tailed)	0.997	0.013	–	0.464	0.007	0.068	0.000	0.000
X_4								
Pearson correlation	0.038	0.084	−0.030	1.000	0.096	0.028	−0.051	0.028
Significance (2-tailed)	0.355	0.041	0.464	–	0.020	0.499	0.216	0.492
X_5								
Pearson Correlation	0.653	0.985	0.110	0.096	1.000	0.115	0.068	0.460
Significance (2-tailed)	0.000	0.000	0.007	0.020	–	0.005	0.101	0.000
X_6								
Pearson correlation	0.029	0.059	0.075	0.028	0.115	1.000	0.053	0.068
Significance (2-tailed)	0.476	0.151	0.068	0.499	0.005	–	0.199	0.240
X_7								
Pearson correlation	−0.015	0.064	0.894	−0.051	0.068	0.053	1.000	0.622
Significance (2-tailed)	0.724	0.119	0.000	0.216	0.101	0.199	–	0.000
X_8								
Pearson correlation	0.172	0.449	0.580	0.028	0.460	0.048	0.622	1.000
Significance (2-tailed)	0.000	0.000	0.000	0.492	0.000	0.240	0.000	–

Table 5 Pearson's correlation analysis for linear dependency between dependent and independent variables

Dependent variable	Independent variables							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Y								
Pearson correlation	0.014	−0.058	−0.843	0.056	−0.061	−0.057	−0.988	−0.589
Significance (2-tailed)	0.736	0.156	0.000	0.172	0.141	0.165	0.000	0.000
Sample size (n)	402	402	402	402	402	402	402	402

46.80 years (reduced by 18.87 years from arsenic pollution-free life expectancy of 65.67 years). So for any kind of public health measures (installation of arsenic treatment units, establishing health care centers, treated water supply through pipe network, road network to facilitate easy access to nearby hospital, establishing

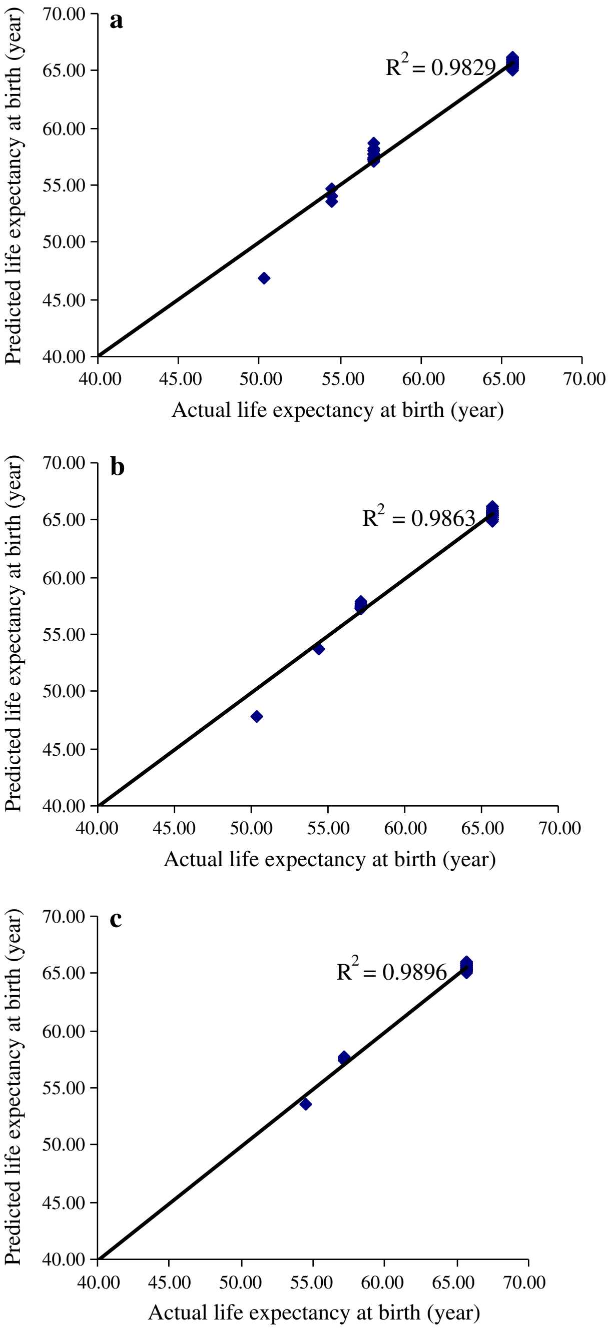

Fig. 5 **a** Relation between actual and observed life expectancy at birth in Domkal block ($n = 87$). **b** Relation between actual and observed life expectancy at birth in Hariharpara block ($n = 62$). **c** Relation between actual and observed life expectancy at birth in Nawda block ($n = 39$)

schools to provide awareness) mouza Jhauberia should be prioritized over Jyotkanai. Usually life expectancy in the study area and in other developing nations is calculated based on crude death rates which do not take environmental degradation factors into account. The model illustrated, by quantifying the impact of arsenic pollution, how important this is for life expectancy.

Conclusions

This paper reported an empirical model to predict arsenic pollution affected life expectancy at birth. Life expectancy is generally calculated using life table techniques, but the empirical model offers an easier calculation of life expectancy, and helps to understand the impact of arsenic pollution at the local scale. The empirical model developed in this study is robust, globally implementable, simple, and needs little data to estimate the impact of arsenic on life expectancy. The model is complementary to the life table technique and will help to reduce the impact of arsenic pollution on life expectancy at the local scale by facilitating better resource planning. This study will help public health engineers and health policy makers to understand the severity of the problem from local to national scale, to facilitate necessary actions, and to mitigate arsenic pollution on a priority basis.

Acknowledgments The authors acknowledge the support of Public Health Engineering Department (PHED) Murshidabad; District Land and Land Reforms Office (DL & LRO) Murshidabad; and Census of India, Directorate of Census Operations West Bengal for providing the relevant data required for the present research work. The authors appreciate Mr. Goutam Roychowdhury, Executive Engineer, Public Health Engineering Department (PHED) Murshidabad for providing the arsenic distribution data of the study area.

References

Abernathy, C. O., Calderon, R. L., & Chappell, W. R. (1997). *Arsenic exposure and health effects*. London: Chapman and Hall Press.

Brown, K. G., Boyle, K. E., Chen, C. W., & Gibb, H. J. (1989). A dose–response analysis of skin cancer from inorganic arsenic in drinking water. *Risk Analysis, 9*, 519–528.

Central Ground Water Board. (1999). *High incidence of arsenic in groundwater in West Bengal: Central Groundwater Board*. New Delhi: Ministry of Water Resources, Government of India.

Chen, C. J., Kuo, T. L., & Wu, M. M. (1988). Arsenic and cancers (letter). *Lancet, 1*, 414–415.

Crimmins, E. M., Saito, Y., & Ingegneri, D. (1997). Trends in disability-free life expectancy in the United States, 1970–1990. *Populant Development Review, 23*, 555–572.

Das, N. G. (2000). *Statistical methods in commerce, accountancy and economics*. Kolkata: Manasi Press.

Elandt-Johnson, R. C., & Johnson, N. L. (1980). *Survival methods and data analysis*. New York: Wiley.

Gaus, I., Webster, R., & Kinniburgh, D. G. (2001). Scales of variation. In Kinniburgh, D. G., Smedley, P. L. (eds.) British Geological Survey Technical Report WC/00/19: Keyworth, pp 161–173.

Lai, D. J., Lee, L. M., & Lee, E. S. (2000). Effects of handicap on life expectancy: The case of China. *Public Health, 114*, 330–335.

Landau, S., & Everitt, B. S. (2004). *A handbook of statistical analyses using SPSS*. Washington, DC: Chapman and Hall/CRC Press, LLC.

McCuen, R. H., & Snyder, W. M. (1986). *Hydrologic modeling: statistical methods and applications*. New Jersey: Prentice-Hall.

Mutafova, M., van de Water, H. P. A., Perenboom, R. J. M., Boshuizen, H. C., & Maleshkov, C. (1997). Health expectancy calculations: a novel approach to studying population health in Bulgaria. *Bulletin of the World Health Organization, 75*, 147–153.

National Research Council (NRC). (Oct. 5, 2004). Arsenic in drinking water. http://www.nap.edu/ 1999.

Newell, C. (1988). *Methods and models in demography*. New York: The Guildford Press.

Nusselder, W. J., van de Velden, K., van Sonsbeek, J. L. A., Lenior, M. E., & van den Bos, G. A. M. (1996). The elimination of selected chronic diseases in a population: The compression and expansion of morbidity. *American Journal of Public Health, 86*, 187–194.

Rogers, A., Rogers, R. G., & Belander, A. (1990). Longer life but worse health? Measurement and dynamics. *The Gerontologist, 30*, 640–647.

Rogers, A., Rogers, R. G., & Branch, L. G. (1989). A multistate analysis of active life expectancy. *Public Health Reports, 104*, 222–226.

Samadder, S. R. (2005). Assessment of risk due to arsenic pollution in drinking water in West Bengal: A GIS approach. PhD thesis 2005. Department of Civil Engineering: Indian Institute of Technology Kharagpur, India.

Samadder, S. R. (2010). Impact of arsenic pollution in drinking water on life expectancy: A GIS approach. *KSCE Journal of Civil Engineering, 14*, 681–691.

Samadder, S. R., & Subbarao, C. (2007). GIS approach of delineation and risk assessment of areas affected by arsenic pollution in drinking water. *Journal of Environment Engineering, 133*, 742–749.

Smith, A. H., Hopenhayn-rich, C., Bates, M. N., Goeden, H. M., Hertz-Picciotto, I., Duggan, H. M., et al. (1992). Cancer risks from arsenic in drinking water. *Environmental Health Perspectives, 97*, 259–267.

Smith, A. H., Lingas, E. O., & Rahman, M. (2000). Contamination of drinking water by arsenic in Bangladesh: A public health emergency. *Bulletin of the World Health Organization, 78*, 1093–1103.

Tseng, W. P., Chu, H. M., How, S. W., Fong, J. M., Lin, C. S., & Yeh, S. (1968). Prevalence of skin cancer in an endemic area of chronic arsenicism in Taiwan. *Journal of the National Cancer Institute, 40*, 453–463.

Valkonen, T., Sihvonen, A. P., & Lahelma, E. (1997). Health expectancy by level of education in Finland. *Social Science & Medicine, 44*, 801–808.

Welch, A. H., & Stollenwerk, K. G. (2003). *Arsenic in groundwater*. Boston: Kluwer Academic Publishers.

Wong, K. M., Strecker, E. W., & Stenstrom, M. K. (1997). GIS to estimate storm–water pollutant mass loadings. *Journal of Environment Engineering, 123*, 737–745.

World Health Organization (WHO). (1984) Guidelines for drinking water quality. Geneva 1: 53.

Wu, M. M., Kuo, T. L., Hwang, Y. H., & Chen, C. J. (1989). Dose–response relation between arsenic well water and mortality from cancer. *American Journal of Epidemiology, 130*, 1123–1132.