Epidemiological and Spatial Association between Arsenic Exposure via Drinking Water and Morbidity and Mortality

Population based studies in rural Bangladesh

NAZMUL SOHEL
Dissertation presented at Uppsala University to be publicly examined in Rosénsalen, Akademiska sjukhuset, 751 85 Uppsala, Ingång 95/96 nbv, Monday, May 17, 2010 at 09:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract

The overall aim of this thesis is to evaluate the risk for increased morbidity and mortality due to long-term arsenic exposure via drinking water by use of epidemiological and spatial approaches in studies performed in Matlab, Bangladesh. A total of 166,934 individuals aged 4 years and above were screened for skin lesions in 2002-2003. Another sample of 115,903 adults aged 15 years or more and a third sample of 26,972 pregnancies in 1991-2000 were considered in a historical cohort and an ecological study, respectively, where risk of adult mortality and spatial clusters of foetal loss and infant death were analysed in relation to arsenic exposure.

More than 70% of the tube-wells in the study area exceeded the threshold for arsenic contamination according to the WHO guideline (10 µg/L). The prevalence of arsenic-induced skin lesions was 3/1000 and men had significantly higher prevalence of skin lesions (SMR 158, 95% CI: 133-188) compared to women. There was an increased risk for death in adulthood due to all non-accidental causes (hazards ratio = 1.16, [95% CI 1.06-1.26]) even at a low level of arsenic contamination (10-49 µg/L). Slightly lower risks were observed for death in cancers (1.44 [1.06-1.95]) and infectious diseases (1.30 [1.13-1.49]) at 50-149 µg/L, but for cardiovascular diseases, it was evident (1.23 [1.01-1.51]) from the level 150-299 µg/L. A dose-response relationship was observed for each of these causes.

We identified high and low risk clusters of foetal loss and infant death that coincided with identified high and low clusters of arsenic exposure.

Water arsenic concentration of the reported main water source was significantly correlated with arsenic concentration in urine, which reflects current arsenic intake from all sources (R²=0.41, p< 0.0001), and the influence of neighbouring water sources was minimal.

The study findings underlines that the ongoing arsenic exposure has resulted in a series of severe public health consequences in Bangladesh that call for reinforcement in the mitigation efforts. Knowledge about the spatial distribution of exposure and health effects may be of value in that process.

Keywords: Arsenic, tube-well, skin lesion, adult mortality, foetal loss, infant death, cohort, spatial model, Bangladesh

Nazmul Sohel, International Maternal and Child Health (IMCH), Akademiska sjukhuset, Uppsala University, SE-75185 Uppsala, Sweden

© Nazmul Sohel 2010

ISSN 1651-6206
urn:nbn:se:uu:diva-121788 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-121788)
Dedication

To my father who is not alive to see this day
Let his soul rest in peace
This thesis is based on the following papers, which are referred to in the text by their Roman numerals I - IV.


Reprints were made with permission from the respective publishers.
Contents

Introduction .................................................................................................................................1
Arsenic contamination worldwide .........................................................................................2
Arsenic contamination: the Bangladesh scenario .................................................................3
  Contamination of soil and surface water ............................................................................4
  Contamination of ground water ..........................................................................................4
  Contamination of food .........................................................................................................4
  Geographical distribution of arsenic exposure .................................................................5
Arsenic contamination and its health effects .........................................................................6
Disease cluster ..........................................................................................................................7
Rationale ..................................................................................................................................8

Aims of the thesis ....................................................................................................................9
Specific objectives ....................................................................................................................9

Materials and Methods ..........................................................................................................10
The study area ..........................................................................................................................10
Study data and population .......................................................................................................11
  Health and Demographic Surveillance System (HDSS) ......................................................12
  Arsenic study in Matlab (AsMat Study) ..............................................................................12
  Geographic Information system (GIS) ................................................................................12
  Study designs and sample sizes ..........................................................................................13
Exposure assessment ..............................................................................................................15
Outcomes ................................................................................................................................15
Covariates and confounders ...................................................................................................16
Statistical analysis ..................................................................................................................16
  Epidemiological studies .....................................................................................................16
  Spatial clusters of arsenic exposure ...................................................................................16
  Spatial clusters of foetal loss and infant death ...................................................................17
  Spatial analysis: validity ....................................................................................................18
Ethical considerations ............................................................................................................18

Results ....................................................................................................................................20
Background data .....................................................................................................................20
Tube-well water concentration ...............................................................................................21
Tube-well water concentration ...............................................................................................22
Individual exposure .................................................................................................................23
Exposure validation ..................................................................................................................24
Prevalence of skin lesions .......................................................................................................25
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>As</td>
<td>Arsenic</td>
</tr>
<tr>
<td>AsMat</td>
<td>Arsenic and health consequences in Matlab</td>
</tr>
<tr>
<td>BGS</td>
<td>British geological survey</td>
</tr>
<tr>
<td>BRAC</td>
<td>Bangladesh Rural Advancement Committee</td>
</tr>
<tr>
<td>BTM</td>
<td>Bangladesh Transverse Mercator</td>
</tr>
<tr>
<td>CHRW</td>
<td>Community health research worker</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic information system</td>
</tr>
<tr>
<td>GPS</td>
<td>Global positioning system</td>
</tr>
<tr>
<td>HDSS</td>
<td>Health and demographic surveillance system</td>
</tr>
<tr>
<td>HG-AAS</td>
<td>Hydride generation atomic absorption spectrophotometer</td>
</tr>
<tr>
<td>HR</td>
<td>Hazard ratios</td>
</tr>
<tr>
<td>ICDDR,B</td>
<td>International Centre for diarrhoeal disease research, Bangladesh</td>
</tr>
<tr>
<td>IDW</td>
<td>Inverse distance weighted</td>
</tr>
<tr>
<td>IQR</td>
<td>Inter-quartile range</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary least square</td>
</tr>
<tr>
<td>RR</td>
<td>Relative risk</td>
</tr>
<tr>
<td>SAR</td>
<td>Simultaneous Autoregressive model</td>
</tr>
<tr>
<td>SES</td>
<td>Socioeconomic survey</td>
</tr>
<tr>
<td>SMR</td>
<td>Standardised morbidity ratio</td>
</tr>
<tr>
<td>VA</td>
<td>Verbal autopsy</td>
</tr>
<tr>
<td>WHO</td>
<td>World health organization</td>
</tr>
</tbody>
</table>
Introduction

Arsenic has a long and evil history, thus its name has become synonymous to poison. In Italy, arsenic poison was used for political assassinations in the 15th and 16th centuries. Some people even believed that Napoleon was poisoned by arsenic-tainted wine served to him while he was in exile [1].

Arsenic is ubiquitous in the environment and presents in more than 200 mineral species, ranks 20th in abundance in the earth's crust, 14th in seawater and 12th in the human body [2, 3]. About one third of the atmospheric change of arsenic is of volcanic origin that is the most important natural sources of arsenic. In nature, arsenic exists in the metallic state in 3 allotropic forms (alpha or yellow, beta or black, gamma or grey) and several ionic forms. It is a heavy metal, primarily used in the production of glass and semiconductors, medicine, pesticides and in timber treatments [3, 4]. Inorganic arsenic of geological origin is also found in ground water that is used for drinking and irrigation purposes in several parts of the world. It often contaminates fruits and vegetables, particularly rice [5].

Arsenic is a naturally-occurring element and the use of drinking-water with elevated arsenic concentration is primarily from natural contamination [6, 7]. The discovery of arsenic in drinking-water in many areas of the world has aroused widespread public health concern. About 100 million people in the world are chronically exposed to inorganic arsenic [6]. The arsenic problem in Bangladesh is perhaps the most devastating by its magnitude of exposure and the number of people affected by it [8, 9]. Chronic exposure to inorganic arsenic through contaminated drinking water causes a large number of adverse health effects including arsenic-induced skin lesions, hypertension and various forms of cancer [4, 6]. Several studies show associations between arsenic in drinking water and cancer [10-20], cardiovascular and peripheral vascular diseases [15, 21-24], diabetes mellitus [25-27] and hypertension [28-31].

However, the arsenic problem of tube well water in Bangladesh was discovered relatively recently. Information on possible adverse health effects in terms of morbidity and mortality in this population is still limited. Knowledge about the spatial distribution of exposure to arsenic through drinking water and clustering of adverse health effects in space caused by arsenic is also scarce.
This project encompasses different study designs to evaluate how arsenic exposure increases the risk for morbidity and mortality and to identify the spatial distribution of arsenic and its association with clusters of morbidity and mortality in space and time.

Figure 1. World map illustrating regions with documented arsenic problems in groundwater (As > 50 μg L⁻¹). [32]

Arsenic contamination worldwide

The level of arsenic concentration in drinking water exceeds the World Health Organization (WHO) guideline maximum contaminant level of 10 μg/L in many parts of the Bengal basin, Mekong basin, Taiwan, Chile, Argentina [33] and in many parts of the USA [34]. It is estimated that about 100 million people in the world, including about 13 million in the U.S, are chronically exposed to inorganic arsenic [6]. This can be the result of human-induced pollution or naturally occurring, as is the case in Bangladesh, where arsenic contained in the sediments dissolves in water aquifers.

Elevated levels of arsenic were reported in either water supplied to the communities or in wells in Argentina, Australia, Bangladesh, Brazil, Canada, China, Ghana, Greece, Hungary, Iceland, India, Japan, Korea, Malaysia, Mexico, Inner Mongolia, Nepal, Romania, Taiwan, Vietnam, Zimbabwe, and the USA [3] (Figure 1). Sources of drinking water vary between and within countries. In Bangladesh and West-Bengal, India, tube-well water is the main source of drinking water. In Northern Chile, the water supply during a certain time period contained arsenic from natural geological sources [35]. Well water was the main source of contamination in Argentina and Taiwan [5].
Arsenic contamination: the Bangladesh scenario

The people of Bangladesh have suffered from water-borne diseases, partly due to the use of polluted surface water from ponds, lakes and rivers [36]. In an effort to provide safe sources of drinking water, shallow tube-wells were installed from the 1970s all over Bangladesh with the assistance from international organizations. In the early 1990s it was discovered that a large proportion of tube-wells was contaminated with arsenic [37]. The arsenic problem in Bangladesh is, perhaps, the most devastating in the world, since half of 6-11 millions hand-pumped tube-wells in the country contains water with an arsenic concentration above 10 μg/L, the threshold for determining arsenic contaminated water according to the WHO [32, 36]. Thus, half of the population in Bangladesh is estimated to be exposed to arsenic contaminated water [8], and one quarter of the inhabitants is exposed to concentrations exceeding 50 μg/L (the Bangladesh national standard) [32, 36, 38, 39].

Figure 2. Map of Bangladesh showing the geological distribution of soil and the distribution of As in ground water [32].
Contamination of soil and surface water

Arsenic is widely distributed in the soil and surface water. Surface water usually contains below 10 μg/litre of arsenic, e.g. in rivers, lakes and ponds [4]. Arsenic concentration in soils ranges from 5 to 3000 mg/kg and arsenic concentrations is very high in the volcanic sediment. The source of arsenic in Argentina, Bangladesh Delta region and Mexico is believed to be volcanic sediment [3, 40, 41]. Himalayan soil contains volcanic sediment with an average concentration of 210 mg/kg of arsenic [40]. Due to geological variation and sedimentation processes in the Delta region, the arsenic-laden soil from the Himalayan [40] has been deposited near the surface and then transported to the depth through cyclic redox conditions [42]. Studies suggest that arsenic may have been transported to the ground water by mechanisms of oxidation, reduction and carbon reduction [5]. Deltaic sands and alluvial sands contain most arsenic in Bangladesh, and therefore the southern and central regions of Bangladesh are most affected (Figure 2). The high arsenic concentrations extend across three major rivers (Padma, Jamuna and Meghna) originated from Himalayan, reflecting the multiple sources of arsenic in that region. There is a good correlation between arsenic in ground water and arsenic in soil indicating a possibility of arsenic formation in soil over time.

Contamination of ground water

Arsenic concentration in Bangladesh aquifers is increased with depth up to certain distance (30-50 meters) and then it is gradually decreased following a bell shape curve [42-44]. This is the average depth of shallow tube-wells used by the Bangladeshis [45]; thus, a major share of arsenic exposure comes from the shallow aquifers. It is believed that the iron results in the high concentration of arsenic in drinking water in that depth [46-48], however, iron is detected in the orange Pleistocene sediment where arsenic concentration is very low [42]. On the other hand, in the gray Holocene aquifer, arsenic concentration in ground water follows the bell shape curve without the presence of iron [42, 48]. Therefore, we may assume that the arsenic-laden sediment may cause the arsenic contamination in ground water. This is consistent with the view that common mechanisms of how arsenic is dissolved from sediment to groundwater operate across the entire basin [5].

Contamination of food

Crops and vegetables, particularly rice, produced in the arsenic contaminated areas may contain arsenic. Contamination of food is more eminent in terms of uptake by plants and toxicity in the flooded areas compared to non-
flooded areas. Plants usually more exposed to arsenic under flooded condition, therefore containing more arsenic. A study conducted in Bangladesh observed a positive correlation between arsenic in ground water and arsenic in rice when comparing 330 samples of Aman and Boro rice throughout the country [49]. This correlation was stronger for Boro rice than for Aman rice. However, a good correlation between arsenic in soil and plants is not always found [50, 51]. Preliminary findings from a nationwide survey of arsenic in soil, crops and irrigation water indicate that the soils in the west and southwest part of Bangladesh contain the highest arsenic concentrations [50].

The staple food in Bangladesh is rice [52], and recent studies indicate that rice contains about 130 μg/Kg dry weight of arsenic, mostly inorganic, with considerable variation [53-55]. Therefore, rice alone can contribute 35-60 μg of arsenic per day when a person consumes 250-500 gram of rice. A similar background exposure level has been observed in our study area [56], as in other parts of Bangladesh [57].

Geographical distribution of arsenic exposure
In 1993, it was discovered that tube-wells were contaminated with arsenic in western Bangladesh [5, 58]. Several regional surveys was initiated 1995-2000 in order to assess the level of ground water arsenic contamination [32, 59]. These surveys demonstrated that the floodplain sedimentary aquifers of all three major rivers of the Bengal Basin (the Ganges, the Brahmaputra and the Meghna Rivers) have elevated concentration of arsenic, mostly in the sediments of Holocene age. However, groundwater arsenic contamination in the Pleistocene and older aquifers of the region are nonexistent or minor [5].

The British Geological Survey (BGS) conducted a nation-wide ground water survey where 3216 water samples were collected based on one sample per 35 km², which were analyzed in the laboratory [32]. Ground water arsenic concentration in the southeast and southwest of Bangladesh and in the Sylhet Basin exceeded 50 μg/L (Figure 3). In our study area, which is adjacent to the lower Meghna River, had an arsenic concentration of 50 μg/L or more in most of the tube wells [59-61]. Lower catchments of all three major rivers (Padma, Jamuna and Meghna) of the Bengal basin had high concentration of arsenic in their ground water [5].

In rural Bangladesh people depend heavily on private or small community owned tube-wells that might be contaminated with arsenic. The level of contamination usually shows substantial variation even within a limited geographic area [62, 63]. Matlab is one of the most severely arsenic affected areas, where 95% of the population use tube-well water for drinking [64]. Therefore, arsenic concentration in their tube-well water may vary
geographically, and accordingly clusters of higher or lower arsenic exposure that may emanate from tube-well water alone or also from food consumed may be found in the area [65].

**Figure 3.** Geographic distribution of arsenic exposure by severity in Bangladesh (District level analysis) [32]

**Arsenic contamination and its health effects**

Arsenic is a known toxic and carcinogenic substance that can be found in ground water [6, 66]. Epidemiological studies show association between arsenic in drinking water and cancers in skin [10], lung [11-15], kidney [16], bladder [14, 15, 17-19] and liver [20]. The earliest symptom of exposure
appears in skin [6, 67], which has a strong tendency to accumulate arsenic bound to the sulphhydryl group in keratin [68, 69]. The skin effects include pigmentation changes and thickening of the outer horny layer of skin (keratosis) [67]. Reportedly, skin lesions appear after 5-10 years of exposure [67] and these skin lesions may turn into more serious and disabling forms including cancer [4, 6, 7, 10, 28, 67, 70]. In Northern Chile, 10% and 5% of all deaths of men and women, respectively, were attributed to arsenic in drinking water in an area where the water supply during a certain time period contained arsenic from natural geological sources [35].

In addition, ingestion of arsenic through drinking-water is associated with several non-cancer diseases, e.g. peripheral vascular disease, hypertension, respiratory diseases, neurological and liver disorders, as well as diabetes mellitus [4, 6, 7]. Ingestion of arsenic is associated with increased risk for cardiovascular diseases [21-23].

Arsenic distribution in space and time and the consequent health effects due to human exposure to the contaminant have been at the epicentre of interest in various studies, e.g. regarding diabetes mellitus [25, 26], hypertension [28], and negative reproductive outcomes [71-73]. However, only a few studies considered space and time in investigating the effect of arsenic exposure in human health. Studies report that arsenic readily passes through human placenta and gave rise to high concentration of arsenic in cord blood as in maternal blood [74, 75]. Therefore, millions of women in the arsenic endemic areas including Bangladesh are exposed to arsenic in their reproductive life, which may have negative effects on their health and on their foetuses and infants.

**Disease cluster**

Morbidity and mortality can be clustered in space due to socioeconomic disparities, cultural differences, variation in food habits, and due to other environmental variations in space. Globally about four million children die every year within the first four weeks of life and a similar number of still births occur, due to socioeconomic differences, education gaps and lack of appropriate health care [76-79].

Arsenic related morbidity and mortality may be clustered in space due to a geographic clustering of arsenic exposure. Arsenic induced skin lesions were found to be clustered in a few villages in Southwest China where arsenic contaminated coal were used for cooking in poor conditions [80]. Several studies have demonstrated that cancers are clustered in space and these clusters may be associated with levels of arsenic exposure in those areas. Blackfoot disease and cancers including liver, nasal cavity, lung, skin,
bladder, kidney, prostate and urinary system were clustered in the southwest and northwest coast of Taiwan [81, 82]. A study from Florida observed 25 clusters of incidence of bladder cancer and these clusters were associated with living in the close proximity to known arsenic contaminated drinking water wells [83]. Another cancer cluster study demonstrated a cluster of childhood leukaemia in Churchill County, Nevada that was suggested to be associated with arsenic exposure [84]

Rationale

It is well established that chronic exposure to arsenic through drinking water increases the risk of cancers, cardiovascular diseases, diabetes, infections, negative reproductive outcomes and many other adverse health outcomes. These outcomes are usually reported to be dose dependent as well as depend on duration of exposure. So far, no study has been conducted to evaluate the association between arsenic exposure and adult mortality in Bangladesh. Several studies have demonstrated that foetal loss and infant death are associated with arsenic exposure, but information on clusters of foetal loss and infant death and their association with arsenic exposure is limited.

Matlab is one of the most arsenic affected areas in Bangladesh where International centre for diarrhoea disease research, Bangladesh (ICDDR,B) has maintained a health and demographic surveillance system for the last four decades. Trained community health research workers (CHRW) collect data on a monthly basis and the databases are being updated routinely. To ensure data quality, supervisors randomly visit some households. ICDDR,B also maintains a geographic information system database from where the locations of households and water sources can be extracted and linked with the demographic database. The linked data sets provide an unique opportunity making valid estimates on risks associated with arsenic exposure and adverse health outcomes in Matlab. It is reported that arsenic exposures varies locally and may form clusters. However, the association between clusters of health outcomes and arsenic exposure has not been evaluated. Such data may help improving the mitigation programs.
Aims of the thesis

The overall aim of this thesis is to evaluate the risk for increased morbidity and mortality due to long-term arsenic exposure via drinking water by use of epidemiological and spatial approaches in studies performed in Matlab, Bangladesh.

Specific objectives

1. To assess the prevalence of arsenic exposure via drinking water and arsenic-induced skin lesions as well as their variation by geographical area, age, gender, and socioeconomic conditions.
2. To analyse whether the long-term arsenic exposure via drinking water has resulted in an increased risk for overall non-accidental adult deaths and/or deaths in cancer, cardiovascular diseases and infections.
3. To evaluate the spatial clustering of arsenic exposure in Matlab, Bangladesh, and to analyse its association to space-time clustering of foetal loss and infant death.
4. To determine whether inclusion of data on the neighbourhood water sources in addition to the main water source would improve the assessment of arsenic exposure through tube-well water, in comparison with the concentration of arsenic metabolites in urine.
Materials and Methods

The study area

The study was conducted in Matlab, a rural Bangladesh sub-district located 53 km south-east of Dhaka. The study area is one of the most arsenic affected parts in the country (Figure 4). The location of Matlab is ideal for the sedimentation process that produces the arsenic laden soil. It is situated near the confluence of the two major rivers, Meghna and Ganges. The area is also a low-lying deltaic plain intersected by numerous canals. During the monsoon, essentially the whole area is flooded except for roads and household compounds that are built on earthen mounds.

Houses in Matlab are usually small, with a single room, mud floor, bamboo walls, and tin roof [85]. Three to six households form a bari, keeping a common courtyard in front of each household and the inhabitants are related through the patrilineal line [65, 86]. Each inhabitant in Matlab area receives a permanent registration number to permanently identify the person within the Health and Demographic Surveillance System (HDSS). Identities of households, bari, and villages are also kept in the database. The individuals’ current location is updated if the person moves from one place to another within the study area, enabling the link of the households to their water sources over time.

Socioeconomic surveys (SES) were conducted in 1974, 1982, 1996 and 2005 providing detailed socioeconomic information including individual educational level and presence of household assets. The majority of the population has very low socioeconomic conditions and is mostly engaged in agricultural production. Usually each bari has at least one tube-well that provides drinking water. There are numerous ponds and canals and the Dhonagoda river in the area, providing an easy access to surface water. The 1996 SES showed that 95% of the population in Matlab used tube-well as the source of drinking water, while water from ponds, river, canals and dug wells were used for other purposes [64]. More than two thirds of the tube-wells in Matlab contain arsenic concentrations exceeding the WHO guideline value of 10 μg/L [63].

The study base was established using the Health and Demographic Surveillance System (HDSS) in Matlab that covers 142 villages and
encompasses a population of 220,000 on 18,386 hectares of land [64]. The HDSS surveillance area is divided into two parts, one with health services provided by the government and the other with ICDDR,B as service provider.

Figure 4. Matlab study area showing ICDDR,B and government service areas.

Study data and population

This project took the advantage of the ongoing HDSS and Geographical Information System (GIS) databases in Matlab, as well as a study entitled “Arsenic and health consequences in Matlab” (AsMat study) conducted in the area in 2002-2003. Outcome and exposure data were obtained from the AsMat study (Paper 1) and covariates were obtained from the HDSS databases. In the cohort study, outcome and covariates were obtained from the HDSS databases and exposure data were taken from the AsMat study (Paper 2). In spatial analysis, outcome and covariates were derived from the HDSS databases, exposure data were taken from AsMat study, and the locations of households, tube-wells and all other spatial data were obtained from the GIS database (Papers 3). In the methodological paper, the data on water arsenic concentration and arsenic concentration in urine were taken from AsMat study and all spatial variables were obtained from the GIS database (Paper 4).
Health and Demographic Surveillance System (HDSS)
Since 1966, the HDSS has been recording all vital and other demographic events in the study area population. Data on birth, migration, marital status, pregnancy outcome and death have been updated based on information collected by the Community Health Research Workers (CHRW) who visits all homes on a monthly basis. CHRWs also on a monthly basis collect information on last menstruation from all married women of childbearing age in the area. If a woman is found not to menstruate they mark the woman as possibly pregnant and record the reported date of last menstrual period (LMP). Pregnancy outcome information is recorded following the definitions used by HDSS. Causes of death in the HDSS area are recorded based on verbal autopsy (VA), a systematic standardised interview [64, 87].

Arsenic study in Matlab (AsMat Study)
This cross-sectional survey of individual water sources over time, presence of arsenic-related skin lesions, and arsenic concentration in water from all tube-wells in the area, was performed from January 2002 to August 2003 [63]. All residents aged 4 years and above were listed from the population database for inclusion in the survey. In total 166,934 individuals were examined for skin lesions and the individuals or the caretakers of the included children were interviewed regarding lifetime water consumption. After completing the interviews, a team of field workers visited the area and collected water samples from all functioning tube-wells (13,286) for analysing arsenic concentration in the water using a field kit [61]. Another sample collected in acidified 24 mL polyethylene tube was marked with the ID number of the tube-wells, transported to Matlab laboratory, and stored at -20°C until the analysis was done. The concentration of arsenic was measured by hydride generation atomic absorption spectrophotometry (Shimudzu AA6800, Shimudzu Corporation, Kyoto, Japan) at ICDDR,B laboratory [63, 88, 89].

Geographic Information system (GIS)
The Geographic Information System (GIS) was integrated with HDSS in 1994, and since then the GIS database has been updated regularly with the spatial information of Matlab baris. Locations of health facilities including hospital and sub centres for the community-based health programs along with spatial features of rivers, canals and sources of drinking/cooking water (tube-wells, ponds, ditches, wells) are also available in the GIS database [90]. During the AsMat project in 2002-2003, a team of surveyor measured the locations of all functioning (13,286) and non-functioning tube-wells (3,215) using global positioning system (GPS), and the GPS data on tube-wells were transformed into Bangladesh Transverse Mercator (BTM), a local
geographic projection system. Each tube-well was assigned the *bari* ID, so that the tube-wells can be linked to the population database. The surveyors identified the locations of 6,459 ponds and 3,369 ditches using the GPS, which were also converted into BTM. However, Matlab residents hardly use water from ditches for drinking or cooking in their daily lives, thus the water from these sources is not included in this analysis. The water samples from all functioning tube-wells were tested for arsenic concentration, which ranged from non detectable (0.5 μg/L) to 3644 μg/L [63].

Study designs and sample sizes

The project includes four research papers. The research design and other related information of those studies are shown in Table 1.

<table>
<thead>
<tr>
<th>Study design</th>
<th>Sample size</th>
<th>Databases</th>
<th>Exposures</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper 1</td>
<td>Cross-sectional 166,334 population</td>
<td>AsMat, HDSS</td>
<td>Water arsenic</td>
<td>Prevalence of skin lesions</td>
</tr>
<tr>
<td>Paper 2</td>
<td>Historical Cohort Study 115,903 population of age ≥15 years</td>
<td>AsMat, HDSS</td>
<td>Water arsenic</td>
<td>Adult death, Foetal loss, infant deaths</td>
</tr>
<tr>
<td>Paper 3</td>
<td>Space-time clustering 27,531 pregnancies</td>
<td>AsMat, HDSS, GIS</td>
<td>Water arsenic</td>
<td>Correlation between U-As and WAs</td>
</tr>
<tr>
<td>Paper 4</td>
<td>Cross-sectional 1307 population</td>
<td>AsMat, HDSS, GIS</td>
<td>Water and urine arsenic</td>
<td>Correlation between U-As and WAs</td>
</tr>
</tbody>
</table>

The population-based survey AsMat included all residents above 4 years of age in Matlab. In total 180,811 individuals were eligible for this survey, and their homes were visited from January 2002 to August 2003. Finally, 166,934 individuals could be interviewed and they were examined for the presence of skin lesions. Arsenic concentration of all functioning tube-wells were measured at ICDDR,B laboratory with a detection limit of 1 μg/L. The arsenic concentration in surface water was arbitrarily set to zero due to its low concentration. Drinking water histories since 1970 were obtained for each individual. Exposure histories were constructed using drinking water histories and data on water arsenic concentrations. Prevalence of skin lesions was measured among men and women (Paper 1).

The historical cohort of individuals 15 years and above was followed from January 1, 1991, until December 31, 2000 (closed cohort). Cases were defined as those who were died from non-accidental causes from January 1, 1991 to December 31, 2000. In total 115,903 individuals were available for analysis in the study. During that period 9,015 individuals had died and 22,488 individuals were lost to follow-up, mainly due to out-migration. The person-time contribution was calculated from January 1, 1991 to date of
death, out migration, or end of study period - whichever came first. Arsenic exposure data were obtained from the survey of arsenic in tube-wells conducted in 2002-2003. The average household water arsenic concentration from 1970 and onwards was used as a measure for individual exposure to arsenic (Paper 2).

In the historical cohort, pregnancies were identified from the data of the HDSS during 1991-2000. In total 29,134 women who also had data on water arsenic concentrations during their pregnancy period were available. Due to missing location in the GIS database and missing socioeconomic information, we excluded 528 and 559 pregnancies respectively. We also excluded 1,075 pregnancies that ended with an induced abortion assuming no association between arsenic exposure and that outcome. Finally 26,672 pregnancies were included in the analysis of identifying purely spatial and spatiotemporal clustering of pregnancy outcomes, and for estimating the relative risk for adverse outcome in those clusters. Using the arsenic concentration in tube-well water and their locations, we detected high and low clusters of arsenic exposure. Purely spatial and spatial-temporal analyses were also performed using circular windows for detecting clusters of foetal loss and infant death. Finally, we superimposed the clusters of foetal loss and infant death on to the clusters of arsenic exposure in order to evaluate association of clustering of foetal loss and infant death with the clustering of arsenic exposure (Paper 3).

Paper 4 is based on analysis of a subset of data from a case-referent study of the risk for skin lesions in relation to arsenic concentration in drinking water. This was part of the AsMat survey conducted in 2002-2003. In total 166,934 individuals were screened for skin lesions, and 504 cases were diagnosed as positive. In the study, 1,830 referents were randomly selected from the Matlab residents. All cases and referents were referred to the health clinics for further investigation. 251 referents refused to go to the clinic, and therefore, spot urine samples were collected from all 504 positive cases and 1,579 referents. These samples were dispensed into 20 mL polyethylene containers, and stored at -20°C. Two urine samples were lost during transportation, thus data on 1,577 referents from 1,307 baris, representing the entire population in Matlab, were available.

The smallest identifiable location available in our GIS database is the bari, and only one individual from each bari was included in the analysis. The individual in a bari was selected randomly if the bari was having more than one member in the data set. Finally 1,307 individuals were available for analysis.

Procedures of measuring arsenic concentration in tube-well water and arsenic concentration in surface water are described earlier. Urine samples
were transported to the Karolinska Institutet, Sweden, for measurement of arsenic metabolites (iAs + MA + DMA) using direct hydride generation atomic absorption spectrophotometer (HG-AAS) [91]. The arsenic concentration was adjusted to the average specific gravity in this population (1.02 g cm\(^{-3}\)), as measured by a refractometer [92]. Finally, linear distance between the household and water source and water arsenic concentrations from tube-wells, river/canals and ponds were used in the analysis (Paper 4).

**Exposure assessment**

A total of 13,286 functional tube-wells were screened for arsenic content. Total water arsenic was determined by hydride generation atomic absorption spectrophotometer (HG-AAS, Shimadzu Model AA-6800) at the ICDDR, B laboratory in Dhaka [63, 89]. Individual level arsenic exposure was determined based on the individuals’ drinking water history and arsenic concentration in all water sources used by the individual from 1970 and onwards (Paper 1). The average arsenic exposure via drinking water for each household was calculated based on the information obtained from the household for a certain year, and this household exposure was used as a proxy for individual exposure in Paper 2. Arsenic exposure for individual pregnancies was based on the concentration of arsenic in the tube-well water used by the women during these pregnancies (Paper 3). Individual level arsenic concentration was compared between water As of reported main source and in urine (Paper 4).

**Outcomes**

Skin lesion may be the first visible manifestation of chronic exposure to arsenic. Skin lesion cases were screened by field workers, identified by physicians and confirmed by experts.

Causes of deaths were determined based on verbal autopsy, a systematic standardised interview [64, 87]. All non accidental deaths were included and those causes were categorized into cancer, cardiovascular diseases, infections, and others.

Early foetal loss or spontaneous abortion is defined as loss of a foetus within 28 weeks of pregnancy excluding induced abortion. Birth of a dead foetus after 28 weeks of gestation was defined as late foetal loss or still-birth. Neonatal, post-neonatal and infant deaths were defined as deaths of live births within 28 days, after 28 days but before 12 months, and before 12 months of age, respectively.
Covariates and confounders

Information of age, gender, assets, and education was obtained from the HDSS. Morbidity and mortality risks in Matlab have reportedly varied with socioeconomic conditions [93], including education [94]. Arsenic exposure has been shown to differ between socioeconomic groups with varying patterns over time [63]. The data on assets were obtained from the household socioeconomic census conducted in 1996 and 2005. Household economic status was measured by constructing a wealth index based on the model for the assets of households in this rural setting [95].

Statistical analysis

Epidemiological studies

Descriptive analysis included calculations of central tendency (means/medians) and variation (percentiles). Average household arsenic exposure was analyzed for cases and populations, for men and women separately, and was tested with ANOVA. Gender, education and asset scores were analysed in relation to exposure and outcome in order to assess possible associations (p≤0.10) using analysis of variance, chi square or Spearman’s correlation coefficients as appropriate for the data being used. Age-adjusted prevalence of skin lesions are given for gender groups as Standardised Morbidity Ratios, using women as reference. The mortality risks in relation to arsenic exposure were estimated by Cox proportional hazards models, adjusting for potential confounders, and adjusted Hazard Ratios (HR) were calculated along with 95% confidence intervals (95% CI). Potential confounding factors that changed the effect estimates by 5% or more were included in the model. We divided average lifetime exposure into quintiles with lowest average arsenic exposure level as the reference. To evaluate sex differentials in the association between arsenic exposure and mortality, we performed stratified analysis by sex. SPSS 16.0 (SPSS Inc, Chicago, IL) software was used for the data analysis.

Spatial clusters of arsenic exposure

The locations of the tube-wells were linked to the population database for mapping spatial patterns of risk for arsenic among Matlab residents. We first created a set of Thiessen polygons around the centroid of tube-wells, i.e. space was allocated to its nearby tube-well point. Local Moran’s I statistic was calculated for detecting high-high (high surrounded by high) and low-low (low surrounded by low) clusters of arsenic exposure as well as for detecting extreme outliers of arsenic exposure and those that were high-low (high surrounded by low) clusters and low-high (low surrounded by high)
clusters [96]. The calculations required spatial proximity matrix (a weight file) that provides information on the configuration and relative location of the polygons. Only the significantly high and low clusters of arsenic exposure were plotted to visualize the cluster and to evaluate its link with the clusters of foetal loss and infant deaths. Calculation, visualization of results, and the use of Local Moran’s I are simple; however, the results near study area boundary are not as reliable as the results from the central part of the study area, because boundary areas exhibit fewer neighbours compared to those in the central part of the study area [97, 98]. Therefore, the clusters near study area boundaries are to be interpreted with care. Geoda (freeware, version 0.9.5i, Geoda Center, Arizona State University, USA) software was used for the analysis.

Spatial clusters of foetal loss and infant death

Spatial scan test implemented in SaTScan® [99] was used to identify unique non-random space-time clusters applying a retrospective space-time permutation model. SaTScan® can detect probable space and space-time locations including multiple clusters in a defined geographic area [100-102]. It uses circles or elliptic and a non-parametric test statistic.

We assumed that the incidence of foetal loss and infant deaths follows a Poisson distribution and under the null hypothesis, and the probability of a foetal loss or infant death in a particular location is proportional to the number of pregnancy outcomes in that location [103]. Using SaTScan®, we estimated the probability that the frequency of foetal loss or infant death at each peak surpasses that expected by chance. We set space and time limitations to 50% that allowed us to scan for both large and small clusters. It took into account the observed number of foetal losses and infant deaths inside and outside the circle when calculating the highest likelihood for each circle. This circle was the most probable cluster, and had a rate that was the least likely to happen by chance alone. The statistical significance of possible clusters was calculated using 999 Monte Carlo simulations [104].

Purely spatial and spatial-temporal analyses were performed using circular windows, and the above analyses were performed unadjusted and also with adjustments for covariates. The output from SaTScan® (Version 7.0.3, Maryland, USA) was imported into SPSS software (Version 16.0, Chicago, Illinois, USA) for evaluating arsenic concentrations between higher and lower risk clusters. A portion of the output was imported in ArcGIS (Version 9.2, California, USA) for mapping significant ($p \leq 0.10$) clusters of higher and lower risk for foetal loss and infant mortality.
Spatial analysis: validity

Simple average and inverse distance weighted (IDW) average arsenic concentration in the five nearest sources (any kind) were used to calculate the proxy for individual arsenic exposure. Pearson’s correlation coefficients were calculated for individual arsenic concentration in drinking water and in urine. Pearson’s correlation coefficients were calculated for individual arsenic concentration obtained using spatial models and arsenic in urine. Arsenic concentration in water and urine was highly skewed, thus this was log transformed before analysis. However, the results are presented using antilog values. We stratified the above analysis by age categories and gender.

Presence of spatial autocorrelation was measured using the Moran’s $I$ statistic, a test statistic that attains values ranged -1 and +1 [105-107]. The statistic is widely used for assessing polygon-based autocorrelation [108, 109]. Multiple linear regression models were used where the dependent variable was the observed arsenic in urine and a vector of independent variables that included arsenic concentration in the most frequently used water source, in the nearest water sources, and a host of socioeconomic and demographic variables. In the first instance the model was run using ordinary least square (OLS). Further, we screened the residuals for the presence of autocorrelation [110]. Such a presence indicates that the assumption of independence in OLS is violated. In this situation, the OLS estimator remains unbiased, but it is no longer efficient, and classical estimation of standard errors will be biased. A remedy to this problem is the use of a model from the family of spatial regression models e.g. the Simultaneous Autoregressive model (SAR). The SAR model can be either “SAR error” model or “SAR lag” model depending upon the value of Lagrange multiplier test statistics developed by Anselin [111]. SPSS 16.0 (SPSS Inc, Chicago, IL) and Geoda (freeware, version 0.9.5i, Geoda Center, Arizona State University, USA) softwares were used for data analysis.

Ethical considerations

The ethical issue in this study is that we identified elevated arsenic levels in the drinking water of the study participants. However, a mitigation program was part of the study and that was initiated in collaboration with Bangladesh Rural Advancement Committee (BRAC), Bangladesh [18]. In total 13,286 functioning tube-wells were identified, and were tested for arsenic content using field kits (Merck, Darmstadt, Germany). These tube-wells were painted in red if arsenic concentration exceeded 50 µg/L (Bangladesh national standard), otherwise painted green. People were advised not to use water from red tube-wells for drinking purpose. High priority was also given
to the households with identified skin lesions and/or pregnant women exposed to arsenic. Individuals with skin lesions and other diseases were given treatment and referral. A series of village information meetings were held before start of the study. People were advised to use treated surface water, rain water or alternative ground water to avoid arsenic contaminated drinking water [58, 61]. Demonstration sites of alternative water sources were constructed by BRAC in the villages. All individuals were informed and were gave consent to participate in the study. An institutional review committee and the ICDDR, B Ethical Review Committee approved the study.
Results

Background data

In the cross-sectional survey a total 166,934 individuals aged 4 years and above (74,408 men and 92,526 women) were screened for arsenic-induced skin lesions and were interviewed about life-time use of drinking water sources. The median age at the time of the study was 28 years for women and 24 years for men.

In the historical cohort study, 115,903 adults aged 15 years and above were included. Of those, 9,015 died during the study period (men=5,013 and women=4,002), 22,488 were lost to follow-up, mainly through out-migration (men=12,975 and women=9,513), and 84,400 were alive at end of study (men=39,111 and women=45,289).

In the spatial analysis, 26,972 pregnancies with different outcomes were included. The average age at pregnancy was 27 years (median: 26, inter-quartile range: 23-31 years).

In the exposure validation study 1,307 individuals aged 4 years or more were randomly selected from the study area. Mean age of the sampled population was 30 years with an equal number males and females.

The average distance from a household to the nearest tube-well, pond and river or canal, the three major water sources used, were 32 (Median 13, Inter-Quartile Range (IQR) 1-49), 87 (Median 59, IQR 35-107) and 175 (Median 126, IQR 59-242) meters, respectively. However, the average distance to the nearest tube-well was 101 (Median 73, IQR 27-145) considering only those tube-wells with arsenic concentration less than 50 μg/L (the Bangladesh national standards). More than 75% of the tube-wells were located within a distance of 50 meter (Figure 5). The median distance to the reported main sources of drinking water was 66 meter (mean distance 123; IQR 15-143 meters).
Figure 5. Access to the different water sources by distance to the source from the household. Cumulative percentages.

The majority of the Matlab population has very low socioeconomic conditions, is mostly engaged in agricultural production and has no or only primary education. Average lifetime arsenic exposure via drinking water was higher for the higher socioeconomic group; however, current exposure is lower for that social stratum. A similar association is observed for the educated and non-educated groups of the population.

**Tube-well water concentration**

Arsenic concentration in tube-well water ranged from <1 to 3644 μg/L. More than 70% of the tube-wells exceeded 10 μg As/L, and more than 60% exceeded 50 μg/L. The distribution was skewed, with a mean of 203 μg/L and median of 167 μg/L. Besides the frequency peak below 1 μg/L, there was a second lower peak around 200-300 μg/L. There was a marked geographical variation (Figure 6). While some high- and low-concentration larger areas may be identified, the local variation was considerable.
Figure 6. Geographic variation in tube-well arsenic concentrations in Matlab, Bangladesh.

Tube-well water concentration

Arsenic concentration in tube-well water ranged from <1 to 3644 μg/L. More than 70% of the tube-wells exceeded 10 μg As/L, and more than 60% exceeded 50 μg/L. The distribution was skewed, with a mean of 203 μg/L and median of 167 μg/L. Besides the frequency peak below 1 μg/L, there was a second lower peak around 200-300 μg/L. There was a marked geographical variation (Figure 6). While some high- and low-concentration larger areas may be identified, the local variation was considerable.
Figure 7. Temporal variation in arsenic concentrations (μg/L) in drinking water of the Matlab population, stratified by household asset score quintiles, where sc 1 is the poorest and sc 5 the richest quintile.

**Individual exposure**

There was a marked increase in arsenic exposure from 1970 to 1997 (Figure 7), and thereafter it started to decline. Cumulative median exposure for women and men was 2335 μg/L and 2132 μg/L respectively. Median exposure for women reached its peak at the age 45-54 years, while for men the highest exposure was observed in the 15-24 age groups. Above 35 years of age the median cumulative arsenic exposure was significantly higher for women than for men ($P < 0.01$).

Exposure to arsenic in drinking water showed a socioeconomic gradient with higher exposure in the higher socio-economic groups defined by household asset score (Figure 7). This was especially so from 1970s to mid-1990s, when the annual exposure to arsenic through drinking water still was increasing. During the last years of the study period the socioeconomic differential in exposure had decreased markedly. People with no education had limited access to tube-wells and those with higher education had switched to other arsenic-free or -low options.
In the historical cohort study, the average arsenic exposure was higher for those with higher education or higher socioeconomic status. Arsenic exposure had a negative association with age. Women had slightly higher average water arsenic concentrations than men. The average level of exposure was 133±118 (median 106 μg/L) and 131±116 μg/L (median 105 μg/L), for deceased and surviving individuals, respectively. However, the mean length of exposure was about the same (12.9±7.7 and 12.6±7.5 years, respectively).

The average concentration of arsenic in the drinking water during pregnancy was 241μg/L (median: 226, inter-quartile range: 38-373). Women with low socioeconomic status and low education had higher frequency of foetal loss and infant death ($\chi^2 = 24.4$, $p < 0.001$) as well as higher exposure to arsenic than those in the higher strata ($F=4.1$, $p=0.003$). Fetal loss and infant death were also higher for women in their first pregnancy or fifth and above pregnancies.

**Exposure validation**

Arsenic in urine was higher than that in the reported main water sources. The geometric mean difference was 92 μg/L (UAs=105, WAs=13) with an intercept of 57 μg/L when plotting arsenic in urine as a function of arsenic in water. Water arsenic concentration from the reported main water sources was significantly correlated with concentration in urine for all ages and for both genders combined ($R^2 = 0.41$, $p<0.0001$). Older women (50 years and above) had lower $R^2$ value than younger individuals. Boys and girls up to ten years of age had similar $R^2$ values (0.54 and 0.53, respectively), but boys had higher difference between urinary arsenic and water arsenic concentrations (88 and 78 μg/L, respectively), as well as higher intercept (53 and 46 μg/L, respectively). Younger children (gender combined, <10 years of age) had higher $R^2$ values than those above 10 years and adults (0.54 and 0.39, respectively).

Different models were tested for estimating individual arsenic exposure, considering simple or weighted averages of five neighbouring water sources (of any kind). Estimated values were positively related with arsenic exposure in urine, but the coefficients of determination were much lower (range of $R^2$: 0.18 - 0.21) than that using the arsenic concentration of the reported main water source ($R^2 = 0.41$).

We observed that arsenic concentration in reported main water source, the average arsenic concentration of the five neighbouring water sources (excluding reported one), and certain age categories were positively related to arsenic concentration in urine (total model $R^2 = 0.46$, intercept=30 μg/L);
while gender had close to significant and education had insignificant effect. We stratified the above analysis by gender and coefficient of determination $R^2$ for men and women were 0.42 and 0.49 respectively. Observed Moran’s $I$ score was 0.167 ($p<0.0001$) indicating possible spatial autocorrelation of the residuals. General model fit improved slightly after spatial adjustment (pseudo $R^2=0.53$, spatial lag model), compared to covariate adjusted regression coefficient ($R^2=0.46$).

**Prevalence of skin lesions**

In total 504 individuals with arsenic-induced skin lesions were identified, yielding a prevalence of 3/1000. Men had higher prevalence than women (standardised morbidity ratio for men 158, 95% CI 133-188, with women as reference). The highest prevalence of skin lesions occurred in the age group 35-44 years for both men and women. There were only two cases below 15 years of age.

Cases of arsenic-related skin lesions were more frequent in the higher socio-economic groups; 37% of cases belonged to the highest asset score quintile (20% in population, $P<0.01$). Similarly, 40% of cases had secondary school education or higher, while this was 17% in the entire study population ($P<0.01$).

**Adult Mortality**

Adult mortality was more frequent in male, in lower educated group, and in individuals with lower asset scores. Deaths due to infections were almost one quarter of all deaths, whereas cardiovascular diseases and cancers constituted 13 and 7 percent among all deaths, respectively. There was a significantly increased risk for non-accidental death with increasing arsenic exposure, even at arsenic concentrations in drinking water of 10-49 μg/L (men and women combined) (hazard ratio = 1.16 [95% confidence interval = 1.06– 1.26]). Significantly increased risks of death in cancer (1.44 [1.06 – 1.95]) and infections (1.30 [1.13–1.49]) were observed at 50-149 μg/L of water arsenic concentration. Similarly, significant association was observed between arsenic exposure and cardiovascular death at exposure levels of 150 μg/L or higher (1.16 [0.96 –1.40]). In all of these analyses there was a clear risk increase with increasing exposure (Figure 6).

We did not observe any gender difference in mortality risk for cardiovascular diseases and infections. Men had slightly higher risk for arsenic-related death in cancers; in contrast, women were slightly more susceptible to death in non-accidental causes. Similar risk estimates were
also observed when stratified the analyses by age group and socioeconomic status.

![Figure 8. Exposure to arsenic in drinking water (μg/L) and risk for adult mortality (Hazard Ratios).](image)

**Arsenic exposure, foetal loss and infant death**

**Spatial analysis of arsenic exposure**

The locations of functioning tube-wells (6,317) used by the women during their pregnancies were used to detect clusters of arsenic exposure. We only considered significant clusters of arsenic exposure for high surrounded by high areas and low surrounded by low areas (high-high and low-low, respectively). We observed high clusters in north and south of Matlab whereas a large low cluster area was observed at the centre of the study area.

**Spatial analysis of foetal loss and infant death**

In the spatial analysis without covariates, two areas were identified significantly different from expected number of cases; one with lower and the other with higher occurrence of foetal loss and infant death than expected under the null hypothesis. A cluster of higher risk was found in south-west of Matlab and one low risk cluster was detected in the central part of Matlab. We observed a relative risk of 1.41 and 0.76 in the high and low risk clusters, respectively. The average arsenic concentration during pregnancy in
the high and low risk clusters was statistically significantly different (p<0.001), 319 µg/L and 174 µg/L, respectively.

The result of the spatial analysis of foetal loss and infant death did not change after inclusion of the covariates (age, parity, education and SES) in the model.

When we superimposed the clusters of foetal loss and infant death on to the clusters of arsenic exposure, the high risk cluster of foetal loss and infant death corresponded to the cluster of high arsenic concentrations in the south-west, and the low risk cluster of foetal loss and infant death coincided to the cluster of low arsenic concentrations. We observed high risk cluster of arsenic exposure in the north of Matlab, but, didn’t observed any significant high risk cluster of foetal loss and infant death. However, the incidence rate of foetal loss and infant death was higher in the north compared to all other Matlab.

Space-time analysis of foetal loss and infant death

We observed a significant high risk cluster of foetal loss or infant death for the period 1992 to 1995 with a relative risk of 1.44. When age, parity, education and SES were added into the model, the previously significant high-risk cluster became non-significant.
Discussion

The results of our studies clearly show that arsenic in drinking water constitutes a serious public health problem in Bangladesh. The prevalence of arsenic-induced skin lesions is a visible reflection of the ongoing exposure to arsenic-contaminated water that affects large groups of the population - 70% of the tube-wells in Matlab had arsenic concentrations exceeding 10 μg/L, the WHO guideline value [8]. The increased risk of foetal loss, infant death and adult mortality is an emerging major threat to public health. For the first time we have shown that the arsenic exposure through drinking water over the past decades is associated with an increased risk of adult death. The arsenic exposure shows a great local variation with "hot spots" of higher exposure levels. We identified high and low risk spatial clusters of incidence of foetal loss and infant death that coincided with the identified high and low clusters of arsenic exposure. The knowledge about this variation in exposure and fatal outcomes may be important when planning mitigation activities. Further, we have shown that water arsenic concentration of the reported main water source was significantly correlated with arsenic concentration in urine (that reflects current arsenic intake from all sources) and the influence of neighbouring water sources was minimal. The analysis also indicated that arsenic exposure via other routes, i.e. food exposure, also plays an important role.

These population-based studies were conducted in an area where ICDDR,B has been maintaining a health and demographic surveillance system (HDSS) for more than four decades that covers the entire population in Matlab, a sub-district of Bangladesh. The data collection of the surveillance system is managed by trained community health research workers (CHRW), who collect demographic information on a monthly basis, and the population database is continuously updated. Socioeconomic surveys are conducted every ten years providing detailed socioeconomic information including household assets, electricity, source of water, sanitation and individual level of education. Spatial information is being updated regularly and integrated with the HDSS in order to identify locations of baris, rivers, canals and sources of drinking and cooking water (tube-wells, ponds, ditches, wells) [90]. To ensure data quality 5% of the demographic events are rechecked by the supervisors. Frequency of non-participation in the surveillance system is very low as well as in the specific studies of this thesis that were embedded in the research framework that Matlab HDSS offers. The surveillance system
covers the entire population living in Matlab, which was used in the large prevalence study that is reported in paper 1. It also serves as an ideal sampling frame for cohort studies, which we used in the historical cohort analysis that is reported in the 2nd paper. The spatial analyses (paper 3 and 4) were enabled by the comprehensive databases including demographic information, arsenic exposure data, as well as spatial data.

We mapped the individual arsenic exposure histories for the entire population in Matlab based on reported water sources used for each year from 1970 to time of survey in 2003, combined with data from analyses of arsenic concentration in all functioning tube-wells in the area. To minimize information bias in the historic tube-well water consumption, the information on drinking water sources was instantly validated using data on source of drinking water from household surveys conducted in 1974, 1982, and 1996. Identification of the water source(s) used by the residents was complicated, because calendar years are not widely used in daily life in rural Bangladesh and the years may be recalled inaccurately. To minimize such effects, we also related the histories of water use to momentous life events and local event calendars.

There had been a steady increase in exposure from the 1970s to the late 1990s, in parallel with the installation of tube-wells. Most tube-wells were installed during 1980s and 1990s [9, 36, 58]. In the latter decade, more than 95% of the population used tube-well water [9]. The arsenic exposure appeared to decrease after 1997, indicating an increasing awareness of the problem and a shift to tube-well water with lower arsenic concentration. Higher socio-economic groups took the lead in shifting to tube-well water sources in the 1970s and 1980s, and seemed again to take the lead to use low or arsenic-free water in recent years. Thus, this association between arsenic exposure and socio-economic group has varied over time [63]. There are usually also associations between socio-economic indicators and mortality outcomes, e.g. asset scores [93] or education [94]. We have considered this when analysing for potential confounders of the results of the studies.

In evaluating the association between arsenic concentration in reported drinking water and arsenic exposure in urine samples, we have found an average intercept between urinary and water arsenic of 57μg/L indicating that the reported water source was not the only source of arsenic exposure. Food and other water sources outside home and neighbourhood, e.g. at work places, may also contribute to the exposure [92, 112]. The staple food in Bangladesh is rice [52] and often uncooked rice contains more than 100 μg/Kg of arsenic that is mostly in the form of inorganic arsenic. This means that a daily intake of 500 g rice would correspond to ingestion of 65 μg of arsenic per day, which is similar (48 μg arsenic per day on the average) to results from other studies in Bangladesh [57]. Similar background exposure
levels were also observed in other studies conducted in Matlab [56]. Subgroup (men, older age) had indications of larger contributions to arsenic exposure from sources other than the reported main water source. This may be explained by differences in daily life patterns and food consumption [52]. Although tube-well water is the main source of drinking water, it may be less frequently used for cooking [64]. In rural Bangladesh, some of the tube-wells contain elevated concentration of iron, which may give a metallic taste and discolour the rice [113, 114]. However, the use of tube-well water for cooking also varies depending on availability of alternative water sources [115]. The reported main source of drinking water could explain 40% of the variability of total arsenic intake, which is lower than anticipated. The agreement between arsenic in drinking water and urinary arsenic was higher among women than among men, which may be explained by the fact that women have restricted mobility due to cultural barriers in this rural context [116, 117] and mainly drink water from the reported sources at their bari.

For a similar reason, there were higher $R^2$ values among young children, who mainly drink water from the reported water sources and from school. Men usually travel for their work, and are more likely to drink water from several sources. During working hours, they largely depend on water from their occupational place. Unfortunately, we did not have information on the water sources used at work. Interestingly, older women show a poor correlation between arsenic in water and urine. The older segments of population (both male and female) often use surface water even for drinking, because they are used to it and they often perceive a metallic taste of tube-well water, especially if there is an elevated iron content [113, 114]. This may explain the low association of arsenic in urine and drinking water for both men and women in the older segment of the population.

Drinking water histories were based on recall of water sources used in the past. We also assumed that arsenic concentration in the tube-well water, as determined in 2002-2003, had remained the same since installation of the tube-wells. Information on temporal variation of arsenic concentration in the ground water is limited. Directly measured individual exposure data would have been desirable. The calculated arsenic exposure obviously cannot take into account any time trends in the historical exposure or likely fluctuation in exposure depending on precipitation. This limitation causes uncertainty about arsenic exposure estimates on the basis of assumption that the current arsenic concentrations were also those of the past. However, in a highly contaminated area there was a fairly stable arsenic concentration for a three-year period [118]. Similarly, the British Geological Survey repeatedly followed randomly selected samples from all over Bangladesh for one and a half year and observed no differences [32]. Further, we followed about 60 tube wells from the AsMat study three times per year over three years without observing any significant trend or major difference over time [73]. Thus, we infer that current information on arsenic concentration in tube-wells is still valid for the current exposure.
wells is a suitable proxy for the previous years. Another potential bias in the adult mortality analysis is the use of household exposure data (and to some extent *bari*) as a proxy for individual level exposure. Since arsenic contamination in tube-well water may vary locally, this could potentially introduce a non-differential bias. However, the methodological exercise comparing individual and household exposure data in a sub-sample of the populations supports the use of the employed strategy.

We identified 504 individuals with arsenic-induced skin lesions in a three-step screening procedure. It included a two-step clinical process of evaluating all skin lesions identified by well trained community health workers in the field, because primary stage cases can easily be overlooked if the skin is not carefully investigated [67, 119]. Water arsenic concentrations were measured after clinical skin examinations to avoid bias in the identification of arsenic-related skin cases. All skin lesions identified by field workers were examined by specially trained physicians in a clinic, who documented and photographed the skin lesions for final confirmation by two independent expert dermatologists who reviewed the photographs. In this way we tried to optimize sensitivity in the initial screening, followed by a two-step effort to exclude the false positive findings.

Information on new pregnancies and pregnancy outcomes were prospectively collected by CHRW during their monthly home visit. By this method the number of very early miscarriages may be underestimated, but miscarriages taking place later or stillbirths will be ascertained. Without a death registry the completeness of case ascertainment may be uncertain [120]. Information on deaths in all ages are collected by the CHRW and recorded in the HDSS, and a group of trained people classify the cause of death based on systematic and standardized interviews [64, 87]. This technique, where causes of death are classified based on information obtained from relatives or associates of the diseased through systematic retrospective questioning is known as verbal autopsy (VA) [87, 121]. Causes of death in the HDSS area are based on VA [64], and a validation against physicians’ reports in Matlab revealed a high specificity of verbal autopsy classification (more than 95%), while the sensitivity for cancer deaths was 85% and cardiovascular deaths varied between 50% and 80% [122]. Similar results have been observed in India and China [123]. However, there are no reasons to believe that misclassification of cause of death is associated with arsenic exposure levels. The study is also stemmed from clear *a priori* hypothesis that increased risk would be found for cancer, cardiovascular diseases, infections and non-accidental adult mortality based on previous studies in Chile, Taiwan, Argentina and USA [6, 7, 124].

As mentioned above we have tried to analyse and consider possible socio-economic confounding in the analyses. A number of studies have indicated
that people with poor nutrition are particularly susceptible to arsenic related health effects [67, 125-127]. This may at least partly be catered for when including socio-economic indicators in the analyses. The possibility of residual confounding cannot be excluded, and nutritional status may also play a role as an effect modifier between arsenic exposure and outcome.

About 60% and 70% of the tube-wells in Matlab were found to have arsenic concentration exceeding 50 μg/L, Bangladesh standard, and 10 μg/L, the WHO guideline value [8] respectively. The arsenic concentration was highly skewed with 9.4% exceeding 500 μg/L. This implies that arsenic concentration in water in Matlab is among the highest in Bangladesh [32], most likely because the study area is located in a place where the Meghna River joins the confluence of the Brahmaputra and Ganges rivers, and the ground is highly affected by the historic sedimentation of arsenic laden soils.

The overall prevalence of skin lesions was 0.3%, which is lower compared to previous studies [128-131]. The main reason for the discrepancy is probably that the previous studies were conducted in small, often selected, study populations. The prevalence of skin lesions was higher among men than women, which is consistent with other studies [67, 129, 131-133], although none of those was designed to evaluate gender differences. The results indicate that young men start using tube-well water earlier in life than women. This may be related to the fact that some women move into Matlab by marriage from other areas. Probably, installation of tube-wells started somewhat earlier in Matlab than in surrounding areas. Still, women have in general higher cumulative exposure than men, except for young people. This may be due to more frequent use of surface water by men working in the field, particularly some years ago when tube-wells were less common. Women spend more time at home where the tube-wells were first installed. In spite of the differences in exposure, the prevalence of skin lesions was higher among men than women, which is consistent with other studies, [67, 129, 131-133]. However, higher susceptibility among women has also been reported [134]. The mechanism behind this gender difference is not clear.

Compared to women, men often have a higher fraction of the monomethylated arsenic metabolite, MMA, in urine [135, 136], which has been associated with increased risk for arsenic-related skin lesions including skin cancer [137, 138]. Other susceptibility factors, possibly involved in the observed gender differences, include total water intake, sun exposure, smoking habits, and genetic. Men are usually more exposed to the sun than women in Matlab, where rice cultivation and fishing are the most common occupations among men. In contrast, women are mainly occupied in domestic work. Since about 70% of adult men in rural Bangladesh smoke, compared to less than 1% of the women, the role of smoking in the observed gender differences need to be evaluated. Both arsenic and smoking are potent inducers of oxidative stress [139-142], and a recent small-scale study
suggests that genetic susceptibility to oxidative stress is associated with elevated risk of developing arsenic-related hyperkeratosis [143]. Also, arsenic skin lesions in Inner Mongolia were related to markers of oxidative DNA damage [144]. However, no association was found between smoking and arsenic-related skin cancer [137], or cutaneous squamous cell carcinoma (SCC) in general [145].

For the first time we have shown that the arsenic exposure via tube-well water in Bangladesh has started to be associated with increased risk of adult mortality. This risk increase was demonstrated already at relatively low levels of arsenic exposure, i.e. 10-49 μg/L, in the drinking water. There was a strong association between arsenic exposure and non-accidental adult mortality. In agreement with arsenic-related increased mortality risk reported from Taiwan [146], we observed 16-36% increased mortality risks associated with arsenic in the adult population. The possible effects on mortality outcomes in Bangladesh may increase even more, because several tube wells were installed in the late 1990s and the health consequences may be obvious after a few decades. In addition, recent studies from Chile suggest that exposure to arsenic very early in life greatly increases the mortality risk in both malignant and non-malignant lung disease later in adulthood [14]. Mortality due to infections showed a strong association with exposure and a clear dose-response relationship. Previous studies have shown elevated risk of chronic respiratory diseases in individuals who drink arsenic contaminated water [14, 147]. We observed statistically significant increased risk of cancer and cardiovascular mortality, although at somewhat higher exposure level than for overall non-accidental deaths. The discrepancy may partly be explained by the fact of rather limited time of exposure (13 years on average) that may not be enough to demonstrate statistically valid associations of arsenic in drinking water with cancer and cardiovascular deaths. The reported latency period for cancer and cardiovascular diseases varies from 20-40 years [6]. In rural Bangladesh, smoking prevalence is fairly high among men, while hardly any woman smokes cigarettes [148]. Various forms of cancer, cardiovascular diseases and some of the infectious disease mortality are known to be associated with cigarette smoking that might have contributed to the outcome [12, 16]. However, data on smoking was not available within this study, which thwarts an adequate analysis of gender differentials in mortality risks. There are few previous epidemiological studies designed to evaluate gender differences in arsenic-related health risks [149]. In the present study, we have not observed any major gender differences in mortality risk, but men are in slightly higher risk for arsenic-related cancers. A previous study in Chile indicated that arsenic-exposed men had higher cancer risk compared to women [35]. Possibly, the fact that men smoke much more than women in both these countries, may contribute to the higher risk among men. We have previously shown that men in the currently studied population were more...
susceptible to arsenic-induced skin lesions in the form of pigmentation changes and hyperkeratosis [88]. This may at least partly be related to the less efficient metabolism of arsenic in men compared to women [150] which is a known risk factor for various forms of arsenic toxicity, including skin and bladder cancer [149].

Several studies have already demonstrated that individual level arsenic exposure may increase the risk of foetal loss, infant death, low birth weight and several other adverse health consequences for the reproductive outcome [73, 151-154]. We have previously shown an association between arsenic exposure during pregnancy and an increased risk of infant mortality, most likely in deaths caused by infections [73]. Possibly, there is an interaction with the prevalent malnutrition, as this is likely to decrease the defence against pro-oxidants, such as arsenic, and to affect arsenic metabolism [75, 147, 155]. Arsenic exposure has also been found to affect immune function in children [156]. A number of mechanistic studies support an effect of arsenic on immune function [157]. Based on these earlier findings it is logical to expect an association between the arsenic exposure (that shows a variation in space) with space clustering of foetal loss and infant mortality. We observed several hotspots of arsenic exposure in our study area from where we may expect excess risk of other adverse health outcomes, particularly for those morbidities triggered even by short-term exposure to arsenic.

Cluster analysis plays an important role in detecting spatial aggregation of disease or fatal outcomes in order to identify underlying environmental factors of these health problems. We have identified clusters of foetal loss and infant death in Matlab, which is more spatial than spatio-temporal. The southern part of Matlab that is close to the Meghna river had a high risk cluster of foetal loss and infant death, while in the central part of Matlab the incidence was low. We also identified high and low risk clusters of arsenic exposure. The southern and northern part of the area had higher arsenic concentration in tube-well water in comparison with the central part of Matlab. When the low and high risk clusters of foetal loss and infant deaths were compared with the layer showing the cluster of high and low arsenic exposure, high and low risk clusters of foetal loss and infant deaths correspond to the high and low exposure areas supporting the possibility that the spatial patterns of foetal loss and infant death are associated with arsenic exposure. Even after adjustment for covariates in the space analysis the average arsenic exposure via drinking water in the high risk cluster remained significantly higher than that in the low risk cluster supporting possible causal link between arsenic exposure and foetal loss and infant death.

Unadjusted space-time analysis yielded a high-risk cluster of foetal loss and infant death. However, this cluster disappeared after adjustment for the
covariates in the model. This cluster may be formed due to the variation in socioeconomic condition and education levels. We may not observe any spatiotemporal patterns in our study area due to relatively stable level of arsenic concentration in tube-well water and use of these tube-wells over this period of time. Sedimentation processes are always very slow processes and it may take hundred of years to change arsenic concentrations in the ground water [42]. We used data for only a ten-year period. If this space-time cluster was formed due to the arsenic exposure, there is a low probability of having different clusters within this ten-year period.

During the screening for arsenic-related skin lesions and arsenic concentrations in drinking water, mitigation activities were initiated in collaboration with the non-governmental organization BRAC [58]. They also provided further advice and practical assistance concerning alternative water sources. Currently pond-sand filters, tube-well filters or rainwater harvesting are being promoted to ensure arsenic free drinking water to the exposed population. These options are considered as short-term alternatives. Although the mitigation activities in the area were initiated by BRAC in 2002-2003 recent studies from Matlab show that the children are still exposed that indicates a failure of the mitigation programs to provide adequate safe water [158]. The long-term solution most likely be piped water supply and optimum use of surface water that has been practiced in other countries, e.g. Taiwan and Chile. If effective mitigation will be further delayed, we may observe an increased incidence of morbidity and mortality like cancer and cardiovascular diseases in the affected areas.
Conclusions

In these population-based studies from the rural delta-land of Bangladesh we have demonstrated that arsenic exposure via tube-well drinking water constitutes a serious public health problem. Almost three quarters of the tube-wells had arsenic concentrations exceeding the WHO guideline value. Other studies have shown that large segments of the population all over Bangladesh are exposed. The prevalence of skin lesions we report is a visible manifestation of the chronic exposure. We have for the first time shown that the ongoing arsenic exposure in Bangladesh has resulted in increased adult mortality risks with a dose-response relationship. Based on earlier observations of increased risk for foetal loss and infant mortality when pregnant women were exposed to arsenic-contaminated water we analysed the spatial pattern of this association. The hot spots of arsenic exposure coincided with clusters of increased risk of foetal loss and infant mortality – a knowledge that may be of value when planning reinforced mitigation activities. We also analysed whether analysis of water from main drinking water source was the best representation of exposure to arsenic. The additional influence of neighbouring water sources was minimal. However, other sources, e.g. the exposure via food were important.

Considering the magnitude of arsenic exposure in Bangladesh, a considerable proportion of the future disease burden may be attributed to arsenic exposure. Hence, public health interventions are urgently needed. For the risk assessment and management it is essential to identify susceptible population groups and hot spots of exposure, in order to use the mitigation resources in a cost effective way given the limited resources of the country.
Acknowledgements

This thesis was written and the research was carried out at International Maternal and Child Health (IMCH), Department of Women’s and Children’s Health, Uppsala University in collaboration with ICDDR,B: Centre for Health and Population Research, Bangladesh. I gratefully acknowledge the financial support provided by the Swedish International Development Authority (Sida).

I would like to thank a number of people who helped me and supported me to write this thesis.

I am indebted to my supervisor, Professor Lars-Åke Persson for his continuous guidance and invaluable suggestions despite his busy schedule. His guidance inspired me to learn epidemiology and public health. I wish to thank him for giving me the opportunity to conduct my PhD program at Uppsala University.

I am extremely grateful to Professor Marie Vahter, my second supervisor, who provided me scientific inputs and critical reviews of the manuscripts. Her suggestions and encouragement during my thesis writing are highly appreciated. I also wish to thank Professor Pavlos Kanaroglou, my third supervisor, for his continuous support and guidance to understand spatial data, to produce maps, and to use spatial tools in epidemiological studies. I would like to thank Dr Eva-Charlotte Ekström and all co-authors who helped me to improve the manuscripts.

I duly acknowledge my indebtedness to Dr. Mahfuzar Rahman and Dr. Peter Kim Streatfield for allowing me to use the relevant data in my thesis. My sincere thanks go to Dr. M Ali for his critical revisions of my manuscript dealing with spatial analysis and providing me necessary suggestions on how to improve the research designs and writing of the manuscripts.

I gratefully acknowledge Cristina Niska Bachelet for her generous administrative support. I also gratefully acknowledge all staff at IMCH for their support.
I would like to thank all staff in Bangladesh who was involved in the AsMat project. I would like to give my special thank to the people of Matlab who participated in the project and provided relevant information.

I would like to say thank you to all PhD students at IMCH, particularly Hanna, Emma, Barni, Mats and Leif for their kind support. I would like to say thank to the Bangladeshi students Anisur, Ashraf, Baitun and Waheed at IMCH who gave me mental support while staying in Sweden without my family.
Lastly, I would like to thank my mother, my beloved wife Kabery and my two children Sabik and Adeeb for their moral support, inspiration and love which made me able to complete this study.

Sarker Mohammad Nazmul Sohel
March 2010
Reference


115. Khan, N.I., D. Bruce, R. Naidu, and G. Owens, Implementation of food frequency questionnaire for the assessment of total dietary...


128. Tondel, M., M. Rahman, A. Magnuson, I.A. Chowdhury, M.H. Faruquee, and S.A. Ahmad, The relationship of arsenic levels in


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 549

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)