Geospatial association between adverse birth outcomes and arsenic in groundwater in New Hampshire, USA

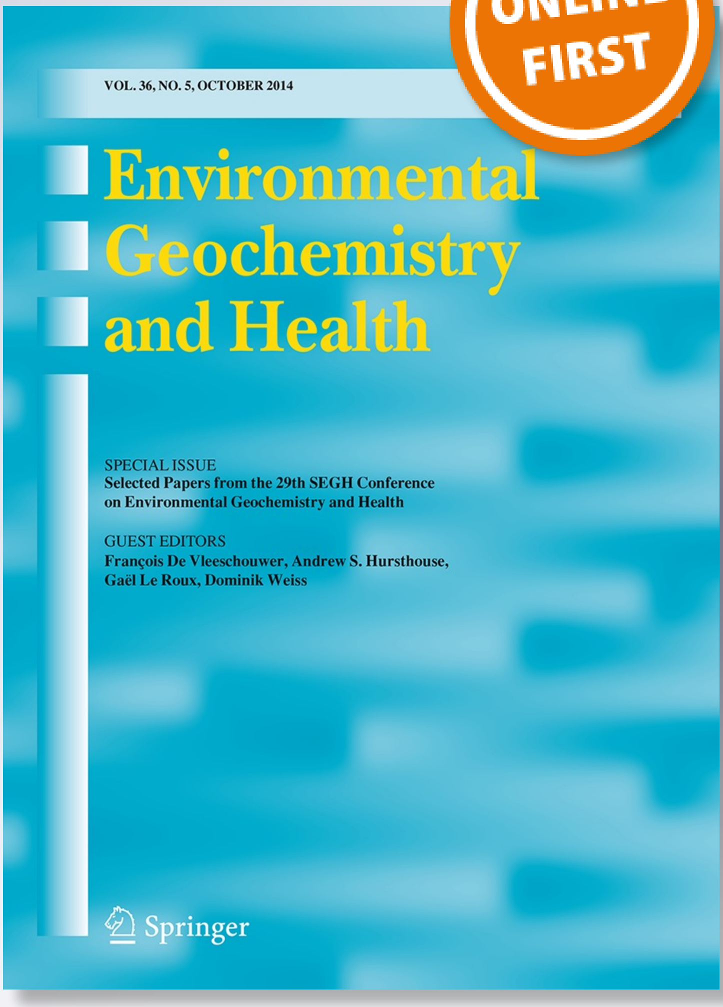
Xun Shi, Joseph D. Ayotte, Akikazu Onda, Stephanie Miller, Judy Rees, Diane Gilbert-Diamond, Tracy Onega, Jiang Gui, Margaret Karagas, et a

Environmental Geochemistry and Health
Official Journal of the Society for Environmental Geochemistry and Health

ISSN 0269-4042

Environ Geochem Health
DOI 10.1007/s10653-014-9651-2

ONLINE FIRST



VOL. 36, NO. 5, OCTOBER 2014

Environmental Geochemistry and Health

SPECIAL ISSUE
Selected Papers from the 29th SEGH Conference on Environmental Geochemistry and Health

GUEST EDITORS
François De Vleeschouwer, Andrew S. Hursthouse, Gaël Le Roux, Dominik Weiss

Springer

Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com”.

Springer

Author's personal copy

Environ Geochem Health
DOI 10.1007/s10653-014-9651-2

ORIGINAL PAPER

Geospatial association between adverse birth outcomes and arsenic in groundwater in New Hampshire, USA

Xun Shi · Joseph D. Ayotte · Akikazu Onda · Stephanie Miller · Judy Rees · Diane Gilbert-Diamond · Tracy Onega · Jiang Gui · Margaret Karagas · John Moeschler

Received: 11 July 2013 / Accepted: 23 September 2014
© Springer Science+Business Media Dordrecht 2014

Abstract There is increasing evidence of the role of arsenic in the etiology of adverse human reproductive outcomes. Because drinking water can be a major source of arsenic to pregnant women, the effect of arsenic exposure through drinking water on human birth may be revealed by a geospatial association between arsenic concentration in groundwater and birth problems, particularly in a region where private wells substantially account for water supply, like New Hampshire, USA. We calculated town-level rates of preterm birth and term low birth weight (term LBW) for New Hampshire, by using data for 1997–2009 stratified by maternal age. We smoothed the rates by using a locally weighted averaging method to increase the statistical stability. The town-level groundwater arsenic probability values are from three GIS data layers generated by the US Geological Survey: probability of local groundwater arsenic concentration >1 μg/L, probability >5 μg/L, and probability >10 μg/L. We calculated Pearson's correlation coefficients (r) between the reproductive outcomes (preterm birth and term LBW) and the arsenic probability values, at both state and county levels. For preterm birth, younger mothers (maternal age <20) have a statewide $r = 0.70$ between the rates smoothed with a threshold = 2,000 births and the town mean arsenic level based on the data of probability >10 μg/L; for older mothers, $r = 0.19$ when the smoothing threshold = 3,500; a majority of county level r values are

---

X. Shi (✉) · A. Onda
Dartmouth College, Hanover, NH, USA
e-mail: xun.shi@dartmouth.edu

A. Onda
e-mail: akikazu.onda@gmail.com

J. D. Ayotte
NH - VT Office, New England Water Science Center,
U.S. Geological Survey, Concord, NH 03301, USA
e-mail: jayotte@usgs.gov

S. Miller · J. Rees · D. Gilbert-Diamond ·
T. Onega · J. Gui · M. Karagas · J. Moeschler
The Geisel School of Medicine at Dartmouth, Hanover,
NH, USA
e-mail: Stephanie.D.Miller2@dartmouth.edu

J. Rees
e-mail: Judith.R.Rees@dartmouth.edu

D. Gilbert-Diamond
e-mail: Diane.Gilbert-Diamond@dartmouth.edu

T. Onega
e-mail: Tracy.L.Onega@dartmouth.edu

J. Gui
e-mail: Jiang.Gui@dartmouth.edu

M. Karagas
e-mail: Margaret.R.Karagas@dartmouth.edu

J. Moeschler
e-mail: John.Moeschler@dartmouth.edu

Published online: 19 October 2014

Author's personal copy

positive based on the arsenic data of probability >10 μg/L. For term LBW, younger mothers (maternal age <25) have a statewide $r = 0.44$ between the rates smoothed with a threshold = 3,500 and town minimum arsenic concentration based on the data of probability >1 μg/L; for older mothers, $r = 0.14$ when the rates are smoothed with a threshold = 1,000 births and also adjusted by town median household income in 1999, and the arsenic values are the town minimum based on probability >10 μg/L. At the county level for younger mothers, positive r values prevail, but for older mothers, it is a mix. For both birth problems, the several most populous counties—with 60–80 % of the state's population and clustering at the southwest corner of the state—are largely consistent in having a positive r across different smoothing thresholds. We found evident spatial associations between the two adverse human reproductive outcomes and groundwater arsenic in New Hampshire, USA. However, the degree of associations and their sensitivity to different representations of arsenic level are variable. Generally, preterm birth has a stronger spatial association with groundwater arsenic than term LBW, suggesting an inconsistency in the impact of arsenic on the two reproductive outcomes. For both outcomes, younger maternal age has stronger spatial associations with groundwater arsenic.

Keywords Preterm birth · Low birth weight · Arsenic · Groundwater · Locally weighted averaging smoothing · New Hampshire

Introduction

Infant mortality can be impacted by adverse reproductive outcomes, and these outcomes can be sensitive to many environmental influences (Wilcox 2001; WHO 2002; Stallones et al. 1992; Braud et al. 2011). Geospatial analysis has been used to study variation in the occurrence of the adverse outcomes (Stallones et al. 1992; Braud et al. 2011; Talbot et al. 2000; Reader 2001; Ozdenerol et al. 2005; Tu et al. 2011) and its possible association with environmental factors, including point-source pollution, such as hazardous waste sites and toxic release sites (Stallones et al. 1992; Braud et al. 2011), pesticide exposure (Xiang et al. 2000), and air pollution (Lin et al. 2004; Slama et al. 2007; Kashima et al. 2011).

In the past 15 years, epidemiologists have been particularly interested in the role of arsenic in the etiology of adverse human reproductive outcomes, especially arsenic exposure from drinking water sourced from groundwater. A geographic concentration of the research is Bangladesh (Ahmad et al. 2001; Milton et al. 2005; Vahter et al. 2006; Kwok et al. 2006; Huyck et al. 2007; Rahman et al. 2009; Tofail et al. 2009; Sohel et al. 2010; Kippler et al. 2012). At the individual level, some studies in Bangladesh provide evidence that arsenic exposure is associated with adverse reproductive outcomes. For example, one study in the MATLAB region identified a significant association in 1,578 mother–infant pairs over the range of urinary arsenic concentrations of <100 μg/L [but not over the entire range measured (6–978 μg/L)]. Over the lower range, a 1.68-g reduction in birth weight was seen for every 1 μg/L increase in urinary arsenic concentration (Rahman et al. 2009). Another study in the Sirajdikhan region analyzed hair, toenail, and drinking water samples collected from 52 pregnant women at multiple time points during pregnancy and from their newborns after birth and suggests that maternal arsenic exposure early in pregnancy negatively affects newborn birth weight (Huyck et al. 2007). However, one study that has examined 2,006 pregnant women from the Faridpur Sadar, MATLAB, and Shahrasti regions who had been chronically exposed to a range of naturally occurring concentrations of arsenic in drinking water only finds small but statistically significant association between arsenic exposure and birth defects and did not see such an association in some other outcomes (Kwok et al. 2006). At the ecological level, a study that examined fetal loss and infant death in the MATLAB region performed geospatial clustering analysis on both reproductive outcomes and arsenic concentration. It finds that the spatial patterns of arsenic concentrations in tube-well water are linked with the adverse pregnancy outcome clusters (Sohel et al. 2010). Besides Bangladesh, an individual-level study in India finds that exposure to high concentrations of arsenic (200 μg/L) during pregnancy was associated with a sixfold increased risk of stillbirth after adjustment for potential confounders, but finds no association between arsenic exposure and spontaneous abortion or overall infant mortality (von Ehrenstein et al. 2006). Ecologic studies in Taiwan and Chile indicate that arsenic endemic areas with drinking

Author's personal copy

water contamination have significantly lower average birth weights compared to non-endemic regions (Yang et al. 2003; Hopenhayn et al. 2003); however, an ecological study in Mongolia did not support an association (Myers et al. 2010). The mechanism through which arsenic influences birth weight is not clear; one of many possible explanations is arsenic-induced impaired glucose tolerance during pregnancy (Ettinger et al. 2009; Andra et al. 2013).

This paper presents a geospatial analysis of associations between groundwater arsenic concentration and two adverse reproductive outcomes, preterm birth and term LBW, in New Hampshire, USA. The novelties of this study include: (1) To our knowledge, geographically this is one of the earliest studies of its kind particularly about a US cohort; (2) this might be one of few studies exploring effect of low concentration of groundwater arsenic on reproductive outcomes; and (3) methodologically, different from most ecological studies of its kind that compare two or a few selected regions, we compare the continuous geographic distributions of arsenic and adverse outcomes. Our conceptual model holds that, if indeed the arsenic in daily drinking water has an effect on human birth, this effect may be revealed by a correlation between the spatial variability of arsenic concentration in groundwater and the spatial variability of adverse reproductive outcomes and that this may be particularly detectable in a region where private wells substantially account for the water supply. In New Hampshire, about 40 % of the population uses private wells as a primary source for drinking water, which is a reason for us to choose it as our study area.

Data

Our choice of preterm birth (gestational period <37 weeks) and LBW (birth weight <2,500 g, e.g., Wilcox 2001) for this study is first based on the availability of data, and is also following suggestions by Wilcox (2001), who provides a powerful discussion of the importance of separating preterm birth and term birth in epidemiological studies. He states that "(a)n exposure that affects fetal growth does not necessarily affect the risk of preterm delivery," and "(c)onversely, a factor that increases the risk of preterm delivery would not necessarily change the average weight of babies delivered at term." He then recommends that when the data of gestational age are available, the preterm birth rate be selected for analysis. For term births, Wilcox recommends to use mean birth weight and simultaneously consider standard deviation (SD), which has been adopted by most studies that are comparing two or a few regions (e.g., Kwok et al. 2006; Yang et al. 2003; Hopenhayn et al. 2003; Myers et al. 2010). However, simultaneously considering mean and SD are difficult when working with many areal units. In such a situation, the LBW rate is convenient and actually to some extent characterizes both mean and SD. Therefore, we choose to use the rate of term LBW (i.e., gestational period ≥37 weeks AND birth weight <2,500 g) as another measurement of adverse reproductive outcome in this study.

Birth data and rate calculation

We obtained birth data from New Hampshire birth certificates for 1997–2009 ($N = 187{,}851$) provided by New Hampshire Department of Health and Human Services (NH DHHS). Each record in the dataset is for an infant and contains information about (1) the infant, including birth date, gestational age, sex, birth weight, plurality, and birth order; and (2) the mother, including age, residential town, and zip code at delivery.

Prior to the analysis, we removed those records of mothers who were not residents of New Hampshire towns, which account for about 1 % of the original records. We then removed those records with a plurality value >1 (i.e., twins and triplets), which account for about 3 % of all records. We also removed those records with apparent invalid or missing values on gestational period, maternal age, and birth weight, which account for less than 1 % of all records. After these processes, a total of 177,995 records remained and were used in the following analyses.

Rate of preterm birth

Among the 177,995 usable records, 12,501 have a gestational period <37 weeks and were identified as preterm births. We stratified the data into detailed categories of maternal age and calculated the preterm birth rate for each category (the upper part of Table 1). The calculation reveals a step at maternal age = 20, so we grouped the detailed categories into two larger categories: maternal age <20 and maternal age ≥20 (the lower part of Table 1). Although the categories of

Author's personal copy

Table 1 Preterm birth ratio by maternal age category, New Hampshire, 1997–2009

	Preterm births	All births	Ratio
Maternal age (years)			
<20	1,034	11,792	0.0877
20–24	2,487	34,272	0.0726
25–29	3,317	49,665	0.0668
30–34	3,400	51,691	0.0658
35–39	1,783	25,368	0.0703
40–44	456	4,975	0.0917
≥45	24	232	0.1034
Two-category stratification			
<20	1,034	11,792	0.0877
≥20	11,467	166,203	0.0690

Table 2 Term low birth weight (LBW) ratio by maternal age category, New Hampshire, 1997–2009

	Term LBW infants	All full-term births	Ratio
Age			
<20	274	10,569	0.0259
20–24	685	31,207	0.0220
25–29	665	47,762	0.0139
30–34	615	47,225	0.0130
35–39	325	22,990	0.0141
40–44	80	4,404	0.0182
≥45	7	198	0.0354
Two-category stratification			
<25	959	41,776	0.0230
≥25	1,692	122,579	0.0138

maternal age ≥40 have greater rates, their relatively small counts for both preterm births and all births may lead to statistical instability and therefore were grouped into the category of maternal age ≥20. We performed the following analyses separately for the two strata to address the influence of maternal age.

Rate of term LBW

The number of usable records of full-term births is 164,335. From these records we identified 2,651 LBW cases (i.e., gestational period ≥37 weeks AND birth weight <2,500 g). Similar to the process with the preterm birth data, we stratified the data into detailed categories of maternal age and calculated the term-LBW rate for each (the upper part of Table 2). The calculation reveals a clear step at maternal age = 24, so we grouped the detailed categories into two larger categories: maternal age <25 and maternal age ≥25 and performed the following analyses separately for these two strata. Similar to the preterm birth data, we grouped the categories of maternal age ≥40 into the category of maternal age ≥25, due to their relatively small counts.

Population and socioeconomic data

The population data used in the rate calculations described above are from the US Census 2010 data (http://www.census.gov/). In this study, however, we did not include the factor of race/ethnicity, because the birth certificates do not contain such information. According to the Census 2010 data, nonwhite females in NH account for 7 % of the female population within the age range 15–49. Thus, we assumed that any race/ethnicity effect is negligible. The risk of birth problems may also be affected by socioeconomic status. We collected town-level income data from the New Hampshire Office of Energy and Planning (www.nh.gov/oep).

Arsenic data

The exposure data used in this study are the modeled probabilities of arsenic occurrence (at thresholds of 1, 5, and 10 μg/L) in private wells that tap groundwater from bedrock aquifers in New Hampshire. This probability of finding arsenic at a location above a given threshold was estimated by using multivariate logistic regression models (“probability models”) developed for New Hampshire (Ayotte et al. 2012). The probability models were developed by using measurements of arsenic from public and private wells as the dependent (or predicted) variable and by using a variety of geologic, geochemical, hydrologic, and anthropogenic data as the independent variable (predictor; Ayotte et al. 2006, 2011, 2012; Flanagan and Ayotte 2011).

Probability models for predicting arsenic concentrations that were greater than or equal to 1, 5, and 10 μg/L in groundwater from bedrock wells were developed in order to produce individual threshold-level probability maps. These three thresholds were

Author's personal copy

chosen because they represent common arsenic reporting levels in groundwater in the State and because they are considered to be possibly relevant concentrations for exposure estimation in terms of potential human health outcomes. Also, the current USEPA maximum contaminant level, the standard for safe drinking water with which public water supplies in the USA must comply, is 10 μg/L. The multivariate logistic regression techniques used to generate the probability estimates are well suited for modeling censored dependent-variable data—data reported as "less than" some laboratory reporting limit—because data that are below reporting limits can be used directly without having to modify or substitute values (Helsel and Hirsch 1992; Helsel 2005; Hosmer and Lemeshow 2000). The well-water arsenic concentration data (dependent data) include censored data that were reported as below laboratory reporting levels (LRLs). The model takes the form:

$$P[y = 1|x] = \frac{e^{(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k)}}{1 + e^{(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k)}} \tag{1}$$

where P is the probability of observing the event, y is an indicator (threshold) variable ("$y = 1$" denoting an event or measurement greater than or equal to a specific value (such as 1, 5, and 10 μg/L), and "$y = 0$" denoting a nonevent or measurement less than a specific threshold), where x_1, x_2,…, x_k are explanatory or independent variables, and where β_0, β_1,…, β_k are unknown parameters (coefficients) to be estimated. The exponential of a parameter, exp (β_i), specifies the proportional increase in the odds of an arsenic concentration being above the modeled threshold per unit increase in the explanatory variable. Threshold values of 1, 5, and 10 μg/L were modeled to identify areas of the State where the probabilities are high for finding low-level (greater than or equal to 1 μg/L) and high-level (greater than or equal to 10 μg/L) arsenic contamination in groundwater. Standard model testing and performance metrics were evaluated and are described in detail elsewhere (Ayotte et al. 2012).

We tested data representing concentrations of arsenic in 1,715 wells (dependent variable) to develop the models, along with more than 250 independent variables, all developed in a geographic information system (GIS), and representing geologic, hydrologic, demographic, and land-use and land-cover features. The final models were dominated by geologic and geochemistry variables but also included variables such as population density, precipitation, groundwater recharge, land use, and proximity to waste sites (Ayotte et al. 2012).

The probability of having arsenic concentrations exceeding 1, 5, and 10 μg/L in groundwater was variable across the state. Generally, high probabilities of arsenic greater than 5 or 10 μg/L were limited to southeastern New Hampshire. However, high probabilities that groundwater from bedrock aquifers would exceed 1 μg/L were widespread across New Hampshire. In fact, nearly half of the State was classified as having at least a 50 % chance of having arsenic greater than or equal to 1 μg/L. High probabilities of arsenic greater than or equal to 5 and 10 μg/L were predicted primarily in the southeastern counties of Merrimack, Strafford, Hillsborough, and Rockingham—the counties that are home to about 75 % of the State's population.

The original USGS arsenic data are in the format of GIS raster layers, with cell size = 30 m. To match the LBW data at the town level, the cell-level data were aggregated to town level by using the Zonal Statistics tool of ArcGIS*, i.e., the average of the values of all the cells falling into a town is used as the representative value of that town (Fig. 1).

Methods

Rate smoothing

To statistically stabilize the rates, we applied a locally weighted average smoothing to the original rates. Locally weighted average methods smooth the rate of an areal unit (in our case, a town) by averaging all the rates of the units in its neighborhood, during which each rate is weighted by its associated background value (in our case, number of births; Kafadar 1994; Waller and Gotway 2004). The specific method we implemented was proposed and justified by Shi et al. (2007), and it is different from conventional locally weighted averaging methods. This method (1) employs a user-specified *background* value rather than a constant geographic distance to define the neighborhood for smoothing, which makes the statistical stability explicit and controllable; (2) generates the neighborhood by creating a buffer around the polygon, rather than about the centroid of the polygon,

Author's personal copy

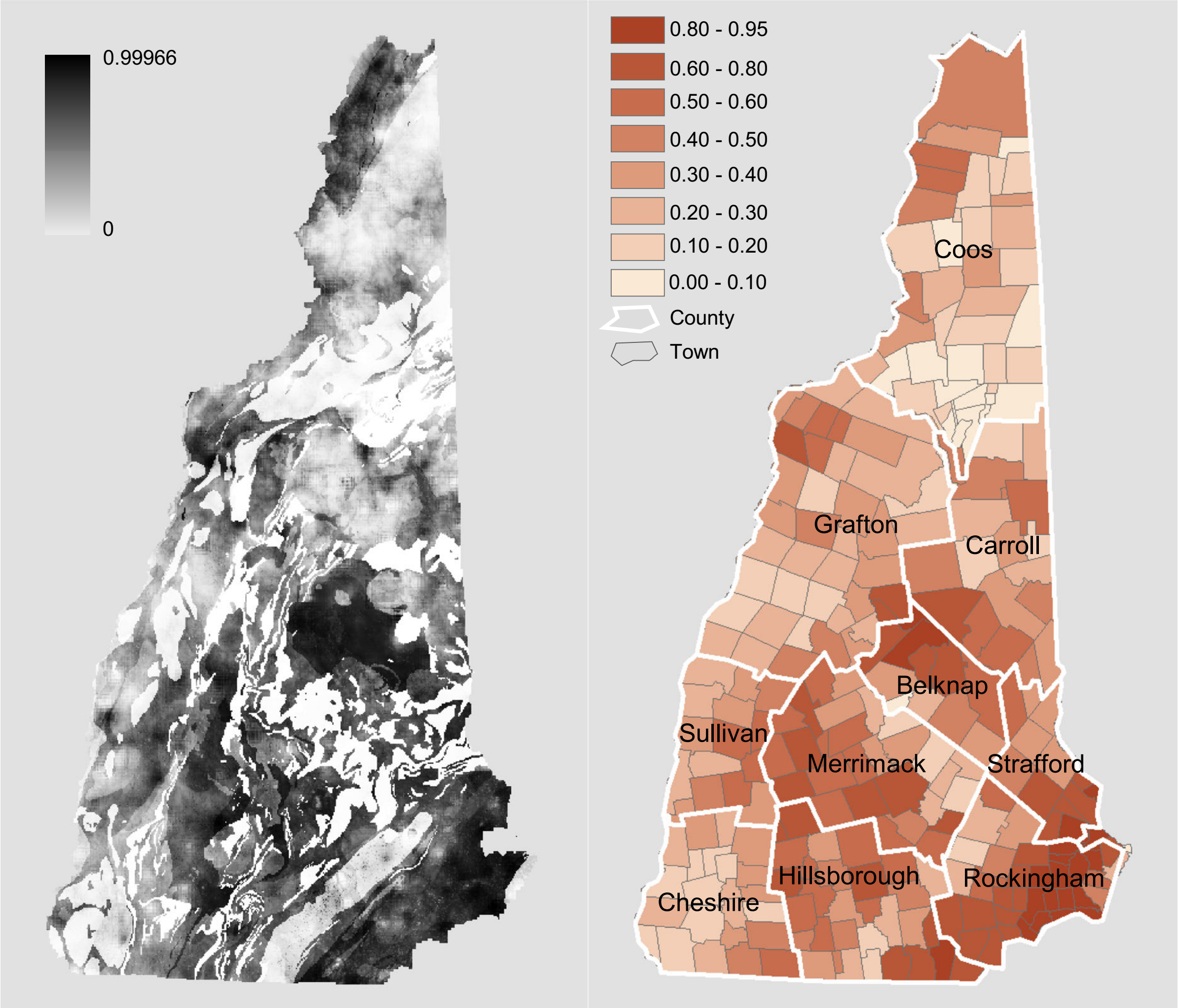


Fig. 1 The USGS modeled probability of groundwater arsenic occurrence in New Hampshire represented as the probability of arsenic concentration exceeding a certain level. The *left* map is the original USGS raster data showing the probability of arsenic concentration >1 µg/L; the *right* map is the town-level mean values calculated from the left map by using the *zonal statistics* tool of ArcGIS*

which takes into account the size and shape of the polygon; and (3) if the neighborhood encloses only part of a polygon, the weight of that polygon will be proportionally determined, allowing a more accurate estimation of contribution of each polygon than an *in-or-out* strategy.

To address the subjectivity in determining the threshold for defining the neighborhood, Shi et al. proposed a strategy that calculates a series of smoothing results using different thresholds. For each of these results, the overall variance of the smoothed rates is calculated and plotted (Shi et al. 2007). It is expected that the variance values become stable as the threshold increases, and the turning point on the plot where the variance value starts to level out is considered as an indication of the optimal threshold. We used this strategy in the current study to identify optimal thresholds.

Correlation calculation

To detect the spatial association between birth problems and groundwater arsenic, we calculated a Pearson's correlation coefficient (r) between the town-level preterm birth rates and the groundwater arsenic levels, and between the term-LBW rates and the

Author's personal copy

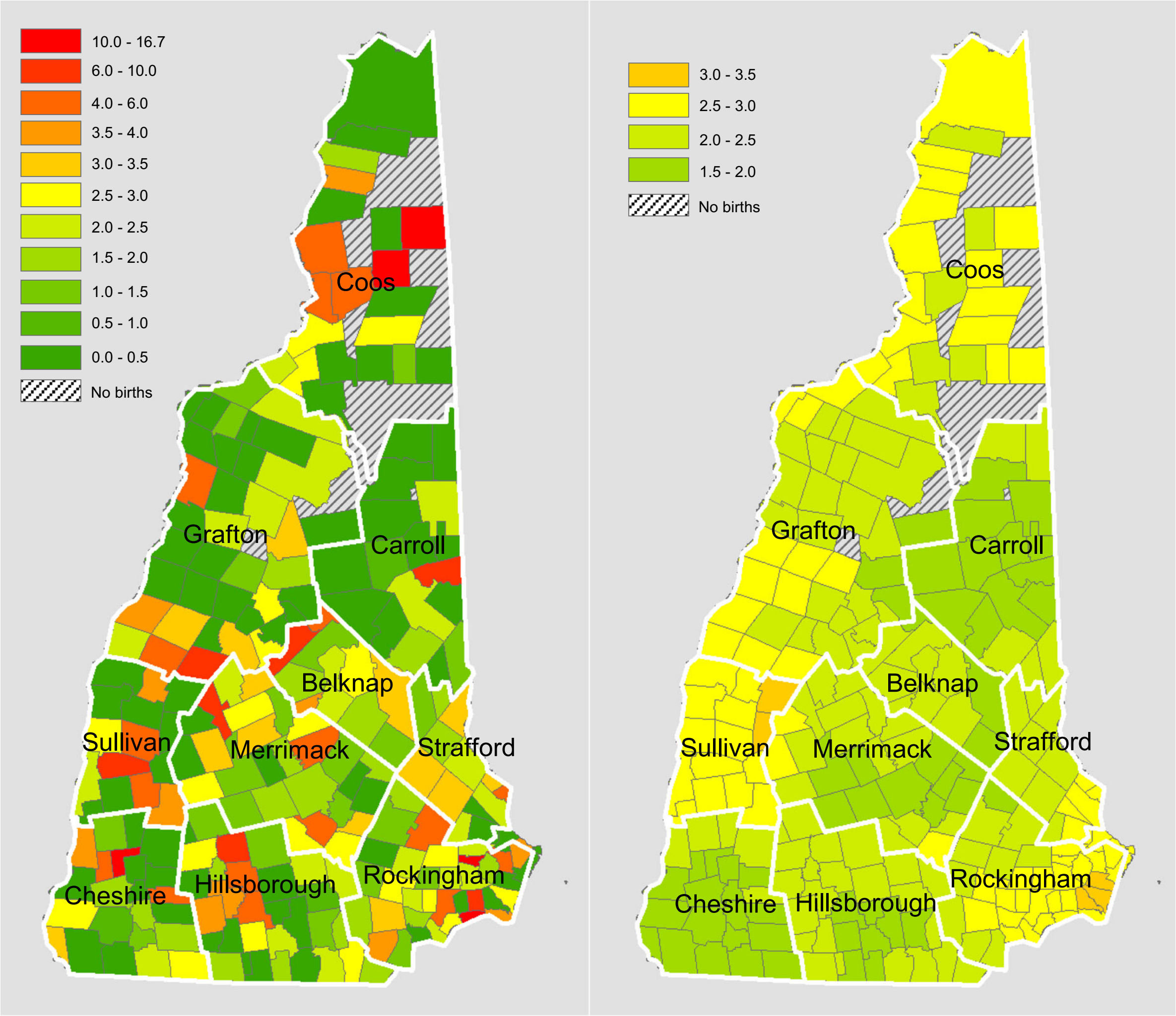


Fig. 2 New Hampshire town-level LBW rates for maternal age <25: the *left* map displays the original rates; the *right* map displays the smoothed rates with smoothing threshold = 2,200 births

groundwater arsenic levels. For both preterm births and term-LBW, the calculation was performed on the original rates and a series of smoothed rates with different thresholds. The calculation was performed for the different maternal age strata separately.

For each stratum, we started the smoothing with a threshold = 100 births, i.e., the buffer around each town polygon was expanded until it enclosed 100 births. With this threshold, those towns with more than 100 births during the study period would not be smoothed. We kept increasing the threshold, by using 100 as the increment, to get smoother results. Figure 2 shows the maps of the original rates and the smoothed rates with a threshold = 2,200 side by side for a visual comparison. These maps are for younger mothers in the term-LBW analysis.

With the map of smoothed rates from a threshold, we calculated *r* between the rates and arsenic level, as well as the variance of rates across the entire state. As an example, Fig. 3 shows the results of these calculations for younger mothers in the term-LBW analysis. Figure 2 indicates that for younger mothers, the correlation between the term-LBW rate and groundwater arsenic almost monotonically increases after the only major drop associated with the starting threshold (100 births). With a threshold of 3,500 births, *r* reaches 0.3. We stopped at 3,500 to avoid over-smoothing. In fact, the variance starts

Author's personal copy

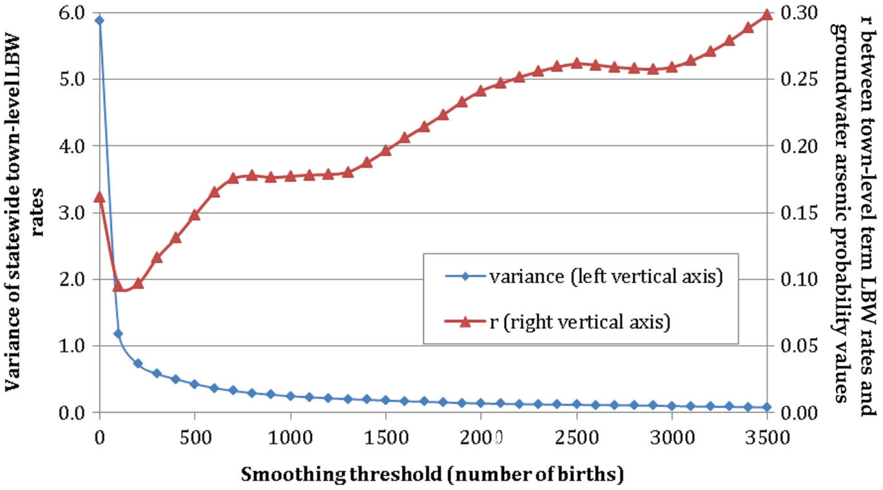


Fig. 3 Variance of New Hampshire town-level LBW rates for maternal age <25 and correlation between the rates and groundwater arsenic probability values against smoothing threshold

to stabilize when the threshold = 2,000 and becomes very small after 2,500.

To eliminate possible impacts of population and income on the risk of birth problems, following Ayotte et al. (2006), we also applied linear regression with the disease rate as the dependent variable and population or income as the independent variable. If the rate was related to population or income, we calculated residuals of the dependent variable and then calculated correlations between the residuals and the arsenic values.

To explore the local variation of the spatial association, we also calculated the correlation coefficient for each of the 10 counties of New Hampshire.

Results

Preterm birth

For the stratum of maternal age <20 (Table 3), statewide, the unsmoothed town-level preterm birth rates have slightly negative *r* values for all three arsenic data layers, as well as for the town household median income value. However, the smoothing changes this situation. Even with a relatively small smoothing threshold of 500 births, *r* values all become positive with considerable magnitudes. Generally, the higher the degree of smoothing, the higher the *r* value. The largest *r* value, 0.70, occurs between the town mean arsenic level based on the data of probability >10 µg/L and the preterm birth rate smoothed with a threshold = 2,000 births (the largest threshold used in this analysis). Among the three arsenic data layers, the ones of probability >5 and >10 µg/L have stronger positive associations with the preterm birth rate than the one of probability >1 µg/L. Among the town minimum, maximum, and mean for the arsenic probability, generally the mean has the highest *r* value, whereas the minimum has the lowest. Unexpectedly, the smoothed rates have fairly considerable positive correlations with the town median household income, i.e., higher rates tend to be associated with higher income values. The adjustment by income consistently lowers down the *r* values, indicating that the income and arsenic may have an association to a certain extent. However, even after the adjustment, the positive association between preterm birth and groundwater arsenic for this group of mothers in New Hampshire is still considerable.

At the county level, generally the *r* values progressively become more positive as the arsenic threshold increases, from probability >1, >5, to >10 µg/L (Fig. 4). For the data of probability >10 µg/L, a majority of *r* values are positive. This progressive variation is most distinct for the five most populous counties (in terms of population density) in New Hampshire, including Hillsborough, Rockingham, Stafford, Merrimack, and Belknap (the five left-most counties in Fig. 4), which account for 78 % of the state's total population, and geographically cluster at the southwest corner of the state that is close to the greater Boston area. For the data of probability >1 µg/L, four of these five counties have dominantly

Author's personal copy

Table 3 Correlation coefficient (*r*) between town-level preterm birth rate for maternal age <20 and groundwater arsenic occurrence in New Hampshire, 1997–2009

	Original Rate	*Smoothed_500*	*Smoothed_1000*	*Smoothed_1500*	*Smoothed_2000*
Prob1					
Town_Min	−0.09	0.25 (0.14)	0.24 (0.14)	0.28 (0.17)	0.31 (0.15)
Town_Max	−0.01	0.37 (0.30)	0.29 (0.22)	0.36 (0.29)	0.43 (0.32)
Town_Mean	−0.08	0.26 (0.15)	0.23 (0.13)	0.30 (0.20)	0.39 (0.23)
Prob5					
Town_Min	−0.06	0.47 (0.37)	0.44 (0.35)	0.50 (0.42)	0.56 (0.42)
Town_Max	−0.13	0.39 (0.26)	0.48 (0.37)	0.57 (0.46)	0.64 (0.47)
Town_Mean	−0.07	0.46 (0.33)	0.48 (0.36)	0.57 (0.45)	0.65 (0.47)
Prob10					
Town_Min	−0.06	0.47 (0.40)	0.46 (0.38)	0.47 (0.40)	0.49 (0.38)
Town_Max	−0.07	0.43 (0.31)	0.41 (0.30)	0.51 (0.41)	0.61 (0.45)
Town_Mean	−0.04	0.57 (0.46)	0.56 (0.47)	0.63 (0.54)	0.70 (0.55)
Median household income	−0.06	0.34	0.30	0.39	0.47

Prob1, *Prob5*, and *Prob10* denote three GIS data layers of modeled groundwater arsenic occurrence, representing probability of arsenic >1, 5, and 10 µg/L, respectively; *Town_Min*, *Town_Max*, and *Town_Mean* denote town-level minimum, maximum, and mean for the modeled arsenic probability values, respectively; *Smoothed_500* etc. denote the smoothed rates; e.g., *smoothed_500* denotes the town-level preterm birth rate smoothed from the *Original Rate* using a threshold that the neighborhood of smoothing must enclose at least 500 births; the value inside the parentheses is *r* adjusted by town-level income (i.e., *r* between the arsenic probability value and the residual to the preterm birth–income linear regression); the bottom line contains the *r* values between the preterm birth rate and the town median household income in 1999; because the original rate has very weak correlation with the income value, the income-adjusted *r* for the original rate is not calculated

negative *r* values, but for the data of probability >10 µg/L, most *r* values of these counties are positive. Another noteworthy finding is that in most cases, the smoothing "helps" increase positiveness.

For the stratum of maternal age ≥20, the *r* values are much smaller compared with their counterparts of the younger mothers (Table 4). While it is hard to claim any significant association based on these *r* values, it seems, however, that the general pattern of them is similar to that of the younger mothers. The probability >5 µg/L and probability >10 µg/L have stronger positive associations with the preterm birth rate than the probability >1 µg/L. In fact, all *r* values based on the 5 and 10 µg/L data are consistently positive, although small. Again, the smoothing generally helps increase positiveness for *r*. The largest *r* value, 0.19, occurs between the town-level mean for probability >10 µg/L and the preterm birth rate smoothed with the largest threshold (3,500 births). However, in this stratum, the preterm birth rate does not appear to have considerable associations with household median income, and therefore, we did not calculate the income-adjusted *r*.

At the county level, the progressive change of *r* values along with the three arsenic data layers is still obvious (Fig. 5). For the data of probability >10 µg/L, a majority of *r* values are positive. This time, the role of smoothing is controversial. For example, for counties of Cheshire and Carroll, smoothing makes *r* values stably and increasingly positive, while for counties of Rockingham, Strafford, and Belknap, smoothing reduces the positiveness.

Term low birth weight

The results of term LBW are generally weaker and less consistent than those of the preterm birth. For the stratum of maternal age <25 years, statewide positive *r* values are dominant across all tests performed (Table 5), but some patterns are different from those of the preterm birth. First, the rank of the three arsenic data layers is reversed and this time the data of probability >1 µg/L have stronger association with the rates than the other two. Second, among the town minimum, maximum, and mean, the minimum consistently has the highest positive *r* values than the other

Author's personal copy

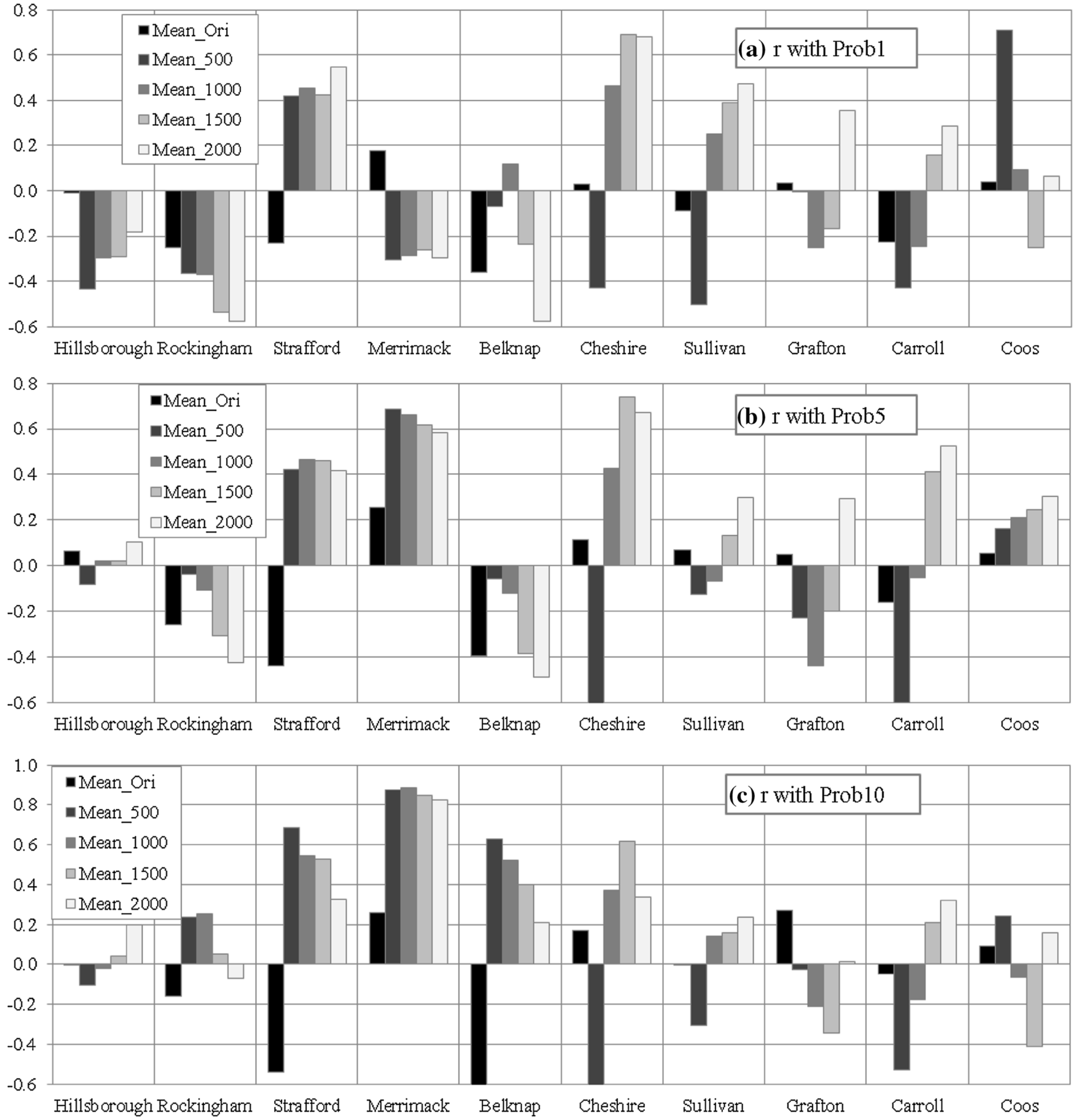


Fig. 4 County-specific correlation coefficient (*r*) between town-level preterm birth rate for maternal age <20 and groundwater arsenic probability values in New Hampshire, 1997–2009. *Notes Prob1*, *Prob5*, and *Prob10* denote three GIS data layers of modeled groundwater arsenic occurrence, representing probability of arsenic >1, 5, and 10 µg/L, respectively; *Mean_Ori* denotes the *r* between the town-level mean probability of arsenic occurrence and the original preterm birth rate; *Mean_500* denotes the *r* between the town-level mean probability of arsenic occurrence and the preterm birth rate smoothed with threshold = 500 births; and so on

two. The smoothing still helps increase the positiveness. The *r* values between the rates and income are fairly small, and therefore, we did not calculate the income-adjusted *r*.

Figure 6 shows the county-specific *r* values between the town minimum arsenic probability value and the term-LBW rate. While statewide more positive *r* is associated with the arsenic probability

Author's personal copy

Table 4 Correlation coefficient (*r*) between town-level preterm birth rate for maternal age ≥20 and groundwater arsenic occurrence in New Hampshire, 1997–2009

	Original Rate	*Smoothed 500*	*Smoothed 1000*	*Smoothed 1500*	*Smoothed 2000*	*Smoothed 2500*	*Smoothed 3000*	*Smoothed 3500*
Prob1								
Town_Min	0.00	−0.04	−0.06	−0.06	−0.05	−0.05	−0.05	−0.05
Town_Max	−0.05	−0.02	0.01	0.04	0.07	0.09	0.11	0.13
Town_Mean	−0.10	−0.10	−0.09	−0.06	−0.02	0.01	0.03	0.04
Prob5								
Town_Min	0.04	0.10	0.09	0.10	0.11	0.12	0.13	0.14
Town_Max	0.01	0.05	0.08	0.10	0.13	0.14	0.15	0.18
Town_Mean	0.01	0.06	0.08	0.10	0.13	0.14	0.16	0.18
Prob10								
Town_Min	0.04	0.06	0.06	0.06	0.07	0.07	0.07	0.07
Town_Max	0.03	0.00	0.01	0.02	0.04	0.05	0.07	0.09
Town_Mean	0.05	0.08	0.10	0.11	0.13	0.15	0.17	0.19
Median household income	−0.12	−0.15	−0.13	−0.11	−0.08	−0.06	−0.04	−0.03

Prob1, *Prob5*, and *Prob10* denote three GIS data layers of modeled groundwater arsenic occurrence, representing probability of arsenic >1, 5, and 10 μg/L, respectively; *Town_Min*, *Town_Max*, and *Town_Mean* denote town-level minimum, maximum, and mean for the modeled arsenic probability values, respectively; *Smoothed_500* etc. denote the smoothed rates; e.g., *smoothed_500* denotes the town-level preterm birth rate smoothed from the *Original Rate* using a threshold that the neighborhood of smoothing must enclose at least 500 births; the bottom line contains the *r* values between the preterm birth rate and the town median household income in 1999; because the preterm birth rate has very weak correlation with the income value, we did not calculate the income-adjusted *r* values

value of 1 μg/L, at the county level, the dominance of positive *r* values is more obvious with the probability value of 10 μg/L. It seems that the inverted results for Merrimack County with the probabilities of 1 and 10 μg/L have caused this controversy. The geographic pattern largely maintains: The three most populous counties (in terms of population density), including Hillsborough, Rockingham, and Strafford, accounting for 62 % of the states' population and clustering near the greater Boston area, generally have positive *r* values across the three arsenic data layers. The smoothing, again, in most cases helps increase the positiveness.

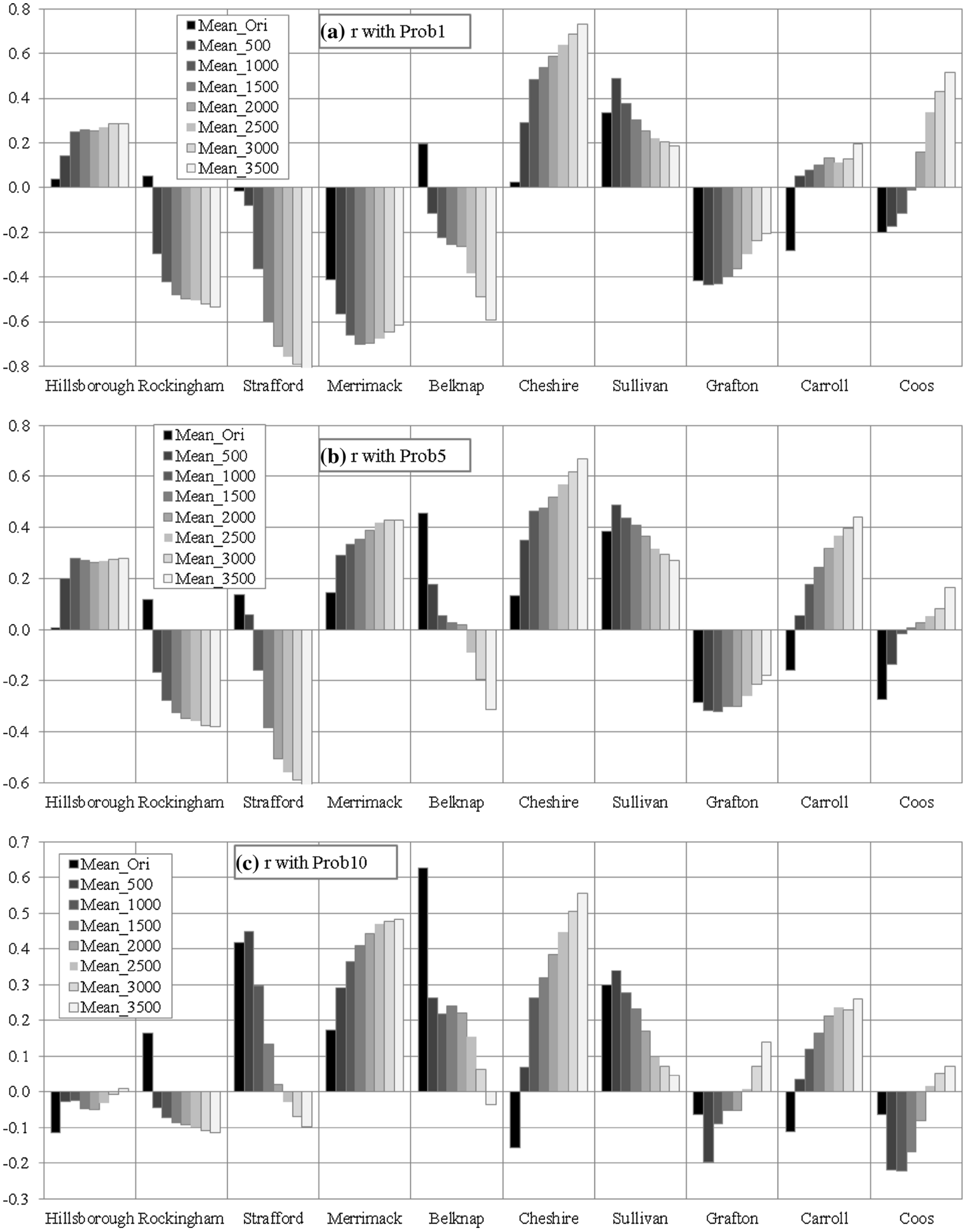
For the stratum of maternal age ≥25 years, statewide *r* is dominated by negative values, a few are fairly considerable (e.g., $r = -0.36$ for the town maximum of probability >10 μg/L and the rates smoothed with a threshold = 2,500 births), although most are very small (Table 6). Along the line from 1, 5, to 10 μg/L, the negativeness generally increases, especially for the town maximum and mean. The effect of smoothing does not have an obvious pattern. The rates, however, have non-negligible negative associations with town median household income (higher rates tend to be associated with lower incomes). The adjustment by the income consistently reduces the negativeness of the *r* values, indicating that income may be a confounder that is overshadowing the effect of groundwater arsenic.

The county-specific correlations for the stratum of maternal age ≥25 years are weak and inconsistent overall. Figure 7 shows the county-specific *r* values between the income-adjusted term-LBW rates and the town minimum arsenic probability. The *r* values are generally small (no matter positive or negative), compared with the stratum of younger mothers. Negative *r* values prevail, in terms of the number of counties, with the data of 1 and 10 μg/L; with the data of 5 μg/L, it is a mix. The most populous county, Hillsborough, maintains to be generally positive across the three arsenic data layers.

Discussion

We found evident spatial associations between two adverse human reproductive outcomes, preterm birth and term LBW, and groundwater arsenic in New

Author's personal copy



Author's personal copy

◀**Fig. 5** County-specific correlation coefficient (*r*) between town-level preterm birth rate for maternal age ≥20 and groundwater arsenic probability values in New Hampshire, 1997–2009. *Notes Prob1*, *Prob5*, and *Prob10* denote three GIS data layers of modeled groundwater arsenic occurrence, representing probability of arsenic >1, 5, and 10 μg/L, respectively; *Mean_Ori* denotes the *r* between the town-level mean probability of arsenic occurrence and the original preterm birth rate; *Mean_500* denotes the *r* between the town-level mean probability of arsenic occurrence and the preterm birth rate smoothed with threshold = 500 births; and so on

Hampshire, USA. However, the properties of these associations vary, in terms of degree of association and sensitivity to different representations of arsenic level. Generally, preterm birth has a stronger spatial association with groundwater arsenic than with the term LBW, suggesting an inconsistency in the impact of arsenic on the two reproductive outcomes, and confirming the necessity to distinguish preterm births and term births in this kind of analysis. For both reproductive outcomes, younger maternal age has stronger spatial associations with groundwater arsenic. In particular, for term LBW, while a positive spatial association between LBW and arsenic level is observed for maternal <25, the association is unclear for maternal age ≥25. However, an initial exploration with town median household income suggests that for the stratum of maternal age ≥25, the effect of groundwater arsenic might have been shadowed by socioeconomic or other factors.

In this study, we treated town median household income as a confounding factor. However, the associations between the reproductive outcome and household income are fairly variable across outcomes and maternal ages. The preterm birth with maternal age <20 has a stronger and unexpected positive association with the household income, and the adjustment by household income consistently reduces the positiveness in the association between the outcome and the arsenic, suggesting that the spatial distribution of household income may co-vary with groundwater arsenic to some extent in NH. The preterm birth with maternal age ≥20 has a slight but negative association with the household income, as well as a much weaker positive association with arsenic, which can be interpreted as that the negative effect of income and the positive effect of arsenic have cancelled each other

Table 5 Correlation coefficient (*r*) between town-level term-LBW rate for maternal age <25 and groundwater arsenic occurrence in New Hampshire, 1997–2009

	Original Rate	*Smoothed 500*	*Smoothed 1000*	*Smoothed 1500*	*Smoothed 2000*	*Smoothed 2500*	*Smoothed 3000*	*Smoothed 3500*
Prob1								
Town_Min	0.17	0.19	0.25	0.30	0.35	0.37	0.40	0.44
Town_Max	0.03	0.00	−0.02	−0.01	0.00	0.01	0.01	0.03
Town_Mean	0.16	0.16	0.19	0.21	0.25	0.26	0.26	0.30
Prob5								
Town_Min	0.21	0.15	0.18	0.23	0.27	0.28	0.29	0.34
Town_Max	0.10	0.02	0.02	0.04	0.08	0.08	0.09	0.14
Town_Mean	0.18	0.10	0.12	0.16	0.21	0.21	0.22	0.27
Prob10								
Town_Min	0.16	0.23	0.25	0.30	0.34	0.36	0.38	0.43
Town_Max	0.04	−0.12	−0.18	−0.18	−0.16	−0.16	−0.15	−0.09
Town_Mean	0.13	0.02	−0.02	0.00	0.04	0.04	0.05	0.11
Median household income	−0.02	−0.03	0.02	0.05	0.10	0.12	0.12	0.18

Prob1, *Prob5*, and *Prob10* denote three GIS data layers of modeled groundwater arsenic occurrence, representing probability of arsenic >1, 5, and 10 μg/L, respectively; *Town_Min*, *Town_Max*, and *Town_Mean* denote town-level minimum, maximum, and mean for the modeled arsenic probability values, respectively; *Smoothed_500* etc. denote the smoothed rates; e.g., *smoothed_500* denotes the town-level preterm birth rate smoothed from the *Original Rate* using a threshold that the neighborhood of smoothing must enclose at least 500 births; the bottom line contains the *r* values between the preterm birth rate and the town median household income in 1999; because the term-LBW rate has very weak correlation with the income value, we did not calculate the income-adjusted *r* values

Author's personal copy

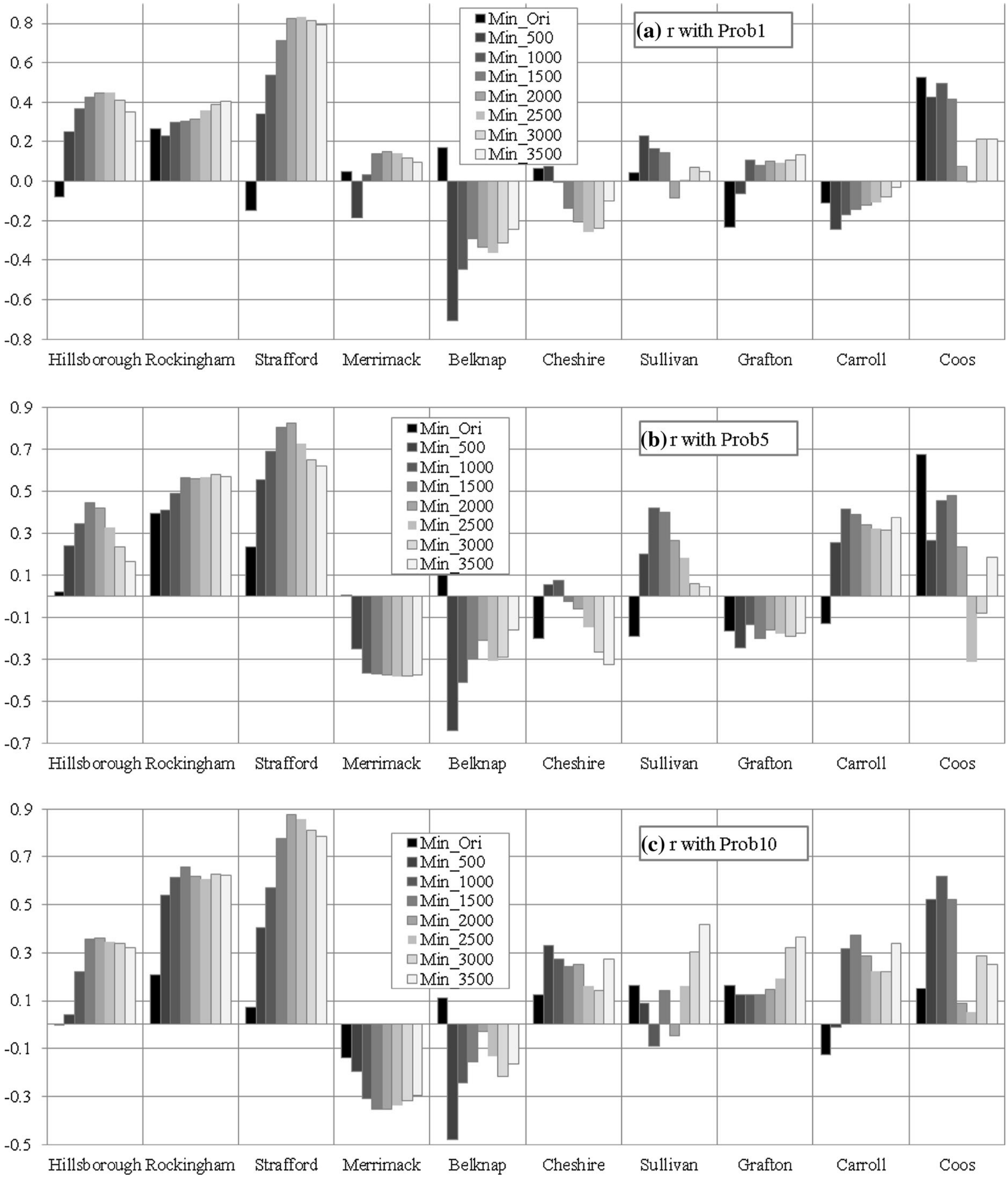


Fig. 6 County-specific correlation coefficient (*r*) between town-level term-LBW rate for maternal age <25 and groundwater arsenic probability values in New Hampshire. *Notes Prob1*, *Prob5*, and *Prob10* denote three GIS data layers of modeled groundwater arsenic occurrence, representing probability of arsenic >1, 5, and 10, respectively; *Min_Ori* denotes the *r* between the town-level minimum probability of arsenic occurrence and the original term-LBW rate; *Minimum_500* denotes the *r* between the town-level mean probability of arsenic occurrence and the term-LBW rate smoothed with threshold = 500 births; and so on

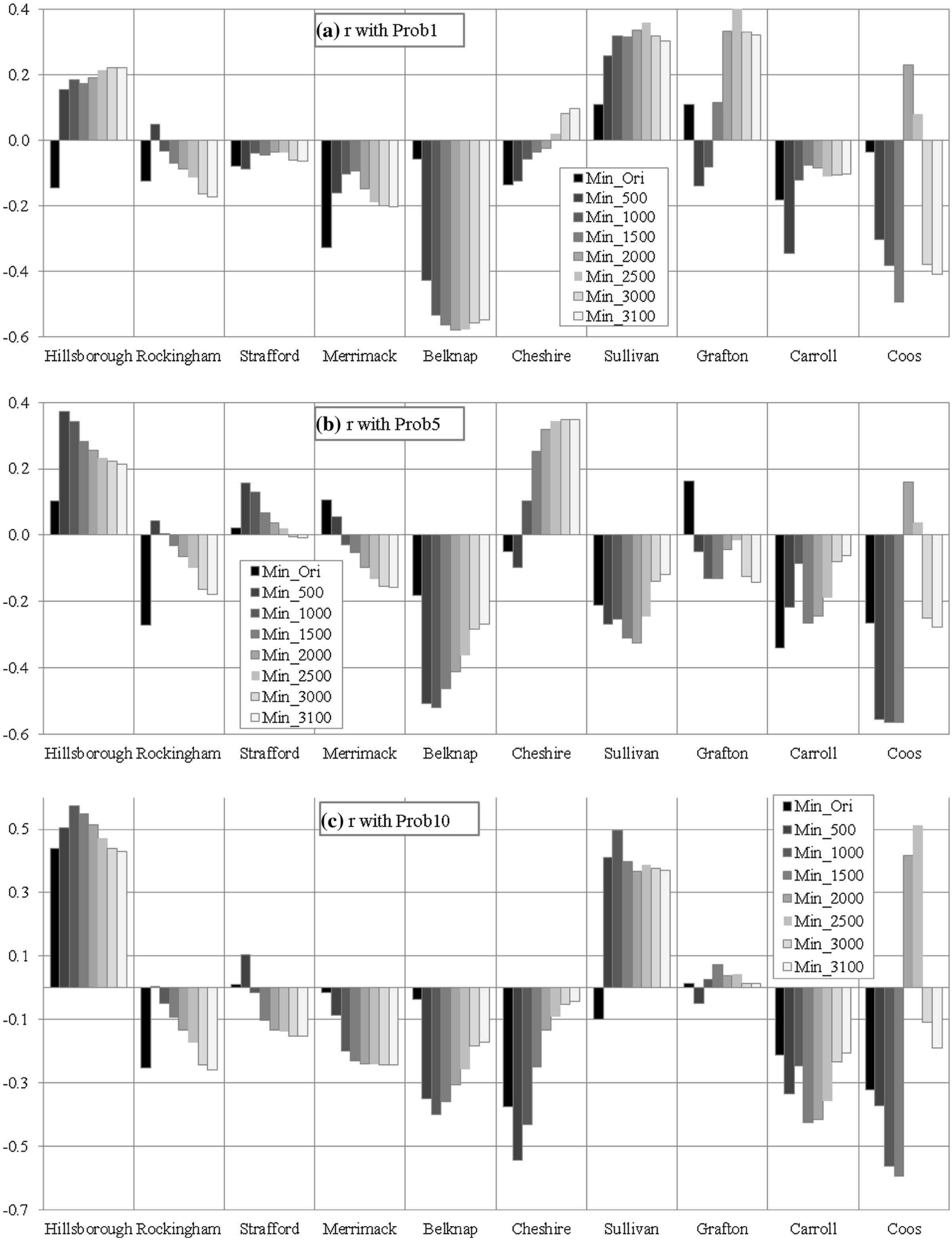
Author's personal copy

Table 6 Correlation coefficient (*r*) between town-level term-LBW rate for maternal age ≥25 and groundwater arsenic occurrence in New Hampshire, 1997–2009

	Original Rate	*Smoothed 500*	*Smoothed 1000*	*Smoothed 1500*	*Smoothed 2000*	*Smoothed 2500*	*Smoothed 3000*	*Smoothed 3500*
Prob1								
Town_Min	−0.03 (0.04)	−0.02 (0.08)	−0.01 (0.12)	−0.03 (0.12)	−0.03 (0.11)	−0.02 (0.11)	−0.01 (0.11)	−0.01 (0.11)
Town_Max	−0.15 (−0.12)	−0.05 (−0.01)	−0.05 (0.00)	−0.06 (0.00)	−0.05 (0.01)	−0.03 (0.02)	−0.03 (0.02)	−0.03 (0.02)
Town_Mean	−0.06 (0.02)	−0.07 (0.02)	−0.10 (0.03)	−0.12 (0.01)	−0.11 (0.02)	−0.09 (0.03)	−0.08 (0.03)	−0.08 (0.03)
Prob5								
Town_Min	−0.03 (0.05)	0.02 (0.13)	0.02 (0.16)	−0.02 (0.13)	−0.05 (0.08)	−0.05 (0.08)	−0.05 (0.08)	−0.05 (0.08)
Town_Max	−0.06 (0.02)	−0.11 (0.00)	−0.12 (0.03)	−0.16 (−0.01)	−0.20 (−0.06)	−0.18 (−0.04)	−0.14 (−0.01)	−0.13 (0.00)
Town_Mean	−0.06 (0.04)	−0.08 (0.04)	−0.08 (0.08)	−0.11 (0.06)	−0.14 (0.03)	−0.12 (0.04)	−0.10 (0.05)	−0.09 (0.06)
Prob10								
Town_Min	−0.01 (0.05)	0.03 (0.12)	0.03 (0.14)	0.00 (0.12)	−0.03 (0.08)	−0.03 (0.08)	−0.03 (0.07)	−0.03 (0.07)
Town_Max	−0.10 (−0.02)	−0.19 (−0.09)	−0.23 (−0.10)	−0.31 (−0.17)	−0.36 (−0.24)	−0.35 (−0.22)	−0.32 (−0.20)	−0.32 (−0.19)
Town_Mean	−0.05 (0.04)	−0.11 (−0.01)	−0.14 (0.00)	−0.20 (−0.05)	−0.24 (−0.10)	−0.24 (−0.10)	−0.22 (−0.09)	−0.22 (−0.09)
Median household income	−0.18	−0.35	−0.37	−0.39	−0.40	−0.37	−0.33	−0.33

Prob1, *Prob5*, and *Prob10* denote three GIS data layers of modeled groundwater arsenic occurrence, representing probability of arsenic >1, 5, and 10 µg/L, respectively; *Town_Min*, *Town_Max*, and *Town_Mean* denote town-level minimum, maximum, and mean for the modeled arsenic probability values, respectively; *Smoothed_500* etc. denote the smoothed rates; e.g., *smoothed_500* denotes the town-level preterm birth rate smoothed from the *Original Rate* using a threshold that the neighborhood of smoothing must enclose at least 500 births; the value inside the parentheses is *r* adjusted by town-level income (i.e., *r* between the arsenic probability value and the residual to the LBW-income linear regression); the bottom line contains the *r* values between the term-LBW rate and the town median household income in 1999

Author's personal copy



Author's personal copy

◀**Fig. 7** County-specific correlation coefficient (*r*) between town-level term-LBW rate for maternal age ≥25 and groundwater arsenic probability values in New Hampshire, 1997–2009. *Notes Prob1*, *Prob5*, and *Prob10* denote three GIS data layers of modeled groundwater arsenic occurrence, representing probability of arsenic >1, 5, and 10 µg/L, respectively; the term-LBW rates used in this figure have been adjusted by the town median household income of 1999; *Min_Ori* denotes the *r* between the town-level minimum probability of arsenic occurrence and the income-adjusted original term-LBW rate; *Min_500* denotes the *r* between the town-level minimum probability of arsenic occurrence and the income-adjusted term-LBW rate smoothed with threshold = 500 births; and so on

to some extent. For term LBW, the correlation between the outcome and household income is noticeable for maternal age ≥25. The adjustment by income reduces the negativeness in the *r* values between LBW rates and arsenic levels. For the stratum of maternal age <25, the correlation between the LBW rate and the income is very weak. While these findings are seemingly variable, they can have a fairly consistent interpretation: Younger mothers, who may have not established a career and/or a stable income, may be less sensitive to the expected negative effect of an economical variable, especially measured at a highly aggregated level, and as a result, the impact of environmental hazards on them might be easier to detect. For older mothers, income might be a stronger independent variable in the equation. This may also indicate that town-level income data may not well represent the economic status of younger mothers.

This study reveals that a smoothing process may have a considerable effect on detection of spatial association. In the analysis of preterm birth and in the analysis of younger mothers with term LBW, it appears that more statistically stable rates (i.e., more smoothing) help reveal potential associations between the reproductive outcomes and the groundwater arsenic. We are aware that for less populous areas such as northern New Hampshire, there is a greater risk of over-smoothing, i.e., the smoothed rates may not correctly reflect the local variability of the disease risk, and in turn, may affect the reliability of the detected association between the disease and environmental exposures.

We find that the spatial associations between the reproductive outcomes and the groundwater arsenic have spatial variation in New Hampshire, and this regional pattern is consistent, as expected, for the two reproductive outcomes we examined. Among the 10 counties of NH, the several most populous counties usually have stronger spatial associations between birth problems and groundwater arsenic than the others. Geographically, these counties are clustered in the southeastern corner of New Hampshire and are more proximate to the greater metropolitan Boston area. There are at least three factors that may have contributed to the stronger correlations they possess: (1) These counties are the most populated, having a majority of the state's population, and, as such, provided a larger sample size with more stable disease rates than in other parts of the state, which may have better reflected the actual influence of groundwater arsenic on reproductive outcomes; (2) some parts of these counties have the highest arsenic probability values in the USGS modeled data, and the high arsenic exposure levels may have increased the detectability of an association between the reproductive outcomes and groundwater arsenic; and (3) compared with the rest of the state, the arsenic probability values vary the most in this region, which also may have facilitated the detection of its spatial association with LBW.

The New Hampshire population uses either public or private water supplies. There is no arsenic regulatory requirement for private water supplies, which results in some having fairly high arsenic concentrations. In contrast, public supplies, by law, must have arsenic monitored and controlled. A limitation of this study is that we did not have precise information to distinguish different water sources in different places. Partially because of this, similar to most geospatial analyses in health studies, the goal of this study has been set to be "exploration." We use geospatial analysis and the best available data to explore if there is a possibility that arsenic in groundwater has an association with adverse reproductive outcomes. To have accurate and precise information about people's source of drinking water, it requires much more extensive and expensive investigations, which is outside the scope of the study presented by this paper. It should also be noted that, as a fairly rural state, New Hampshire's 40 % population using private wells disproportionally occupy a much larger geographic area than the other 60 % of the population. Nevertheless, it will be of interest to make the distinction between private and public water supplies in future data collection and analyses.

Author's personal copy

In this study, we used three GIS data layers of modeled probability of arsenic exceeding certain concentration, including 1, 5, and 10 µg/L. While the results based on the three data layers are generally consistent, the distinction is noticeable. For the preterm birth analysis, the one with the highest bar, probability >10 µg/L, tend to bring about strongest positiveness. For the term-LBW analysis, the one with the lowest bar, probability >1 µg/L, sometimes is more "positive" than the others. What is also noteworthy is the different "performances" of the three representations of town-level arsenic: minimum, maximum, and mean. For the preterm birth analysis, the town mean provides the most positive r values, and for the term-LBW analysis, the town minimum slightly outperforms the other two. A general lesson learned from these findings is that in geospatial analysis for environmental health studies, a thorough consideration and exploration of different data, representations, and parameter settings is necessary.

In principle, the analysis should take into account all confounding factors, subject to availability of data. Maternal age is the only confounding factor we have data at the individual level. Race/ethnicity, a known confounding factor, was not taken into account of this study, due to lack of data of this variable at the individual level. The New Hampshire population is over 95 % Caucasian; hence the impact of this data limitation is minimal. We analyzed the town-level income data as our best-possible effort so far to address socioeconomic factors. Nutritional factors influence biomarker concentrations of arsenic and could potentially be confounders. This is a limitation of our analysis, but we are limited by data in this study, and will certainly take them into account whenever data allow.

Whereas the correlations reported in this paper appear to indicate a relation between the two reproductive outcomes and groundwater arsenic concentration, it does not lead to the conclusion that arsenic from bedrock wells is the true cause of the relation. It is possible that other correlates or combinations of factors that follow a similar pattern to that of arsenic is responsible for the relation we observed. Future investigations of these and other relations may provide additional insight into understanding the effect of exposure to arsenic from private wells in the region.

Finally, we consider it is worth emphasizing the fact that in this study, we found a relation to probabilities of exceeding low-to-moderate arsenic concentration. The finding is novel also because very little is known about adverse reproductive outcomes and arsenic exposure in a US population. It has international value because there are so many parts of the world with low-to-moderate concentrations of arsenic in drinking water that are not generally considered as a health concern.

Acknowledgments Acknowledgements: This study is supported by USNIH grants P20RO18787, P20ES018175, and USEPA grant RD83459901.The authors thank the anonymous reviewers for their insightful comments and suggestions. *ArcGIS is a product name of the Environmental System Research Institute, Redland, California, US. Any use of trade, firm, or product names in this paper is for descriptive purposes only and does not imply endorsement by the US Government.

References

Ahmad, S. A., Sayed, M. H., Barua, S., Khan, M. H., Faruquee, M. H., & Jalil, A. (2001). Arsenic in drinking water and pregnancy outcomes. *Environ Mental Health Perspectives, 109*, 629–631.

Andra, S. S., Makris, K. C., Christophi, C. A., & Ettinger, A. S. (2013). Delineating the degree of association between biomarkers of arsenic exposure and type-2 diabetes mellitus. *International Journal of Hygiene and Environmental Health, 216*, 35–49.

Ayotte, J. D., Cahillane, M., Hayes, L., & Robinson, K. W. (2012). Estimated probability of arsenic in groundwater from bedrock aquifers in New Hampshire, 2011. *U.S. Geological Survey Scientific Investigations Report*, 2012–5156. http://pubs.usgs.gov/sir/2012/5156/.

Ayotte, J. D., Nolan, B. T., Nuckols, J. R., Cantor, K. P., Robinson, G. R., Baris, D., et al. (2006). Modeling the probability of arsenic in groundwater in New England as a tool for exposure assessment. *Environmental Science and Technology, 40*(11), 3578–3585.

Ayotte, J. D., Szabo, Z., Focazio, M. J., & Eberts, S. M. (2011). Effects of human-induced alteration of groundwater flow on concentrations of naturally-occurring trace elements at water-supply wells. *Applied Geochemistry, 26*(5), 747–762.

Braud, T., Nouer, S., & Lamar, K. (2011). Residential proximity to toxic release sites and the implications for low birth weight and premature delivery. *Journal of Environmental Health, 73*(6), 8–13.

Ettinger, A. S., Zota, A. R., Amarasiriwardena, C. J., Hopkins, M. R., Schwartz, J., Hu, H., et al. (2009). Maternal arsenic exposure and impaired glucose tolerance during pregnancy. *Environmental Health Perspectives, 117*, 1059–1064.

Flanagan, S. F., Ayotte, J. D., & Robinson, Jr. G. R. (2011). Quality of water from crystalline rock aquifers in New England, New Jersey, and New York, 1995–2007. *U.S. Geological Survey Scientific Investigations Report*, 2011–5220. http://pubs.usgs.gov/sir/2011/5220/.

Author's personal copy

Helsel, D. R. (2005). *Nondetects and data analysis: Statistics for censored environmental data* (1st ed.). New York, NY: Wiley.

Helsel, D. R., & Hirsch, R. M. (1992). *Statistical methods in water resources*. New York, NY: Elsevier Science Company Inc.

Hopenhayn, C., Ferreccio, C., Browning, S. R., Huang, B., Peralta, C., Gibb, H., et al. (2003). Arsenic exposure from drinking water and birth weight. *Epidemiology, 14*, 593–602.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York, NY: Wiley.

Huyck, K. L., Kile, M. L., Mahiuddin, G., Quamruzzaman, Q., Rahman, M., Breton, C. V., et al. (2007). Maternal arsenic exposure associated with low birth weight in Bangladesh. *Journal of Occupational and Environmental Medicine/American College of Occupational and Environmental Medicine, 49*, 1097–1104.

Kafadar, K. (1994). Choosing among two-dimensional smoothers in practice. *Computational Statistics & Data Analysis, 18*, 419–439.

Kashima, S., Naruse, H., Yorifuji, T., Ohki, S., Murakoshi, T., Takao, S., et al. (2011). Residential proximity to heavy traffic and birth weight in Shizuoka, Japan. *Environmental Research, 111*, 377–387.

Kippler, M., Wagatsuma, Y., Rahman, A., Nermell, B., Persson, L. A., Raqib, R., et al. (2012). Environmental exposure to arsenic and cadmium during pregnancy and fetal size: A longitudinal study in rural Bangladesh. *Reproductive Toxicology, 34*, 504–511.

Kwok, R. K., Kaufmann, R. B., & Jakariya, M. (2006). Arsenic in drinking-water and reproductive health outcomes: A study of participants in the Bangladesh integrated nutrition programme. *Journal of Health Population Nutrition, 24*(2), 190–205.

Lin, C.-M., Li, C.-Y., Yang, G.-Y., & Ma, I.-F. (2004). Association between maternal exposure to elevated ambient sulfur dioxide during pregnancy and term low birth weight. *Environmental Research, 96*, 41–50.

Milton, A. H., Smith, W., Rahman, B., Hasan, Z., Kulsum, U., Dear, K., et al. (2005). Chronic arsenic exposure and adverse pregnancy outcomes in Bangladesh. *Epidemiology, 16*(1), 82–86.

Myers, S. L., Lobdell, D. T., Liu, Z., Xia, Y., Ren, H., Li, Y., et al. (2010). Maternal drinking water arsenic exposure and perinatal outcomes in inner Mongolia, China. *Journal of Epidemiology and Community Health, 64*, 325–329.

Ozdenerol, E., Williams, B. L., Kang, S. Y., & Magsumbol, M. S. (2005). Comparison of spatial scan statistic and spatial filtering in estimating low birth weight clusters. *International Journal of Health Geographics, 4*, 19.

Rahman, A., Vahter, M., Smith, A. H., Nermell, B., Yunus, M., & El Arifeen, S. (2009). Arsenic exposure during pregnancy and size at birth: A prospective cohort study in Bangladesh. *American Journal of Epidemiology, 169*, 304–312.

Reader, S. (2001). Detecting and analyzing clusters of low-birth weight incidence using exploratory spatial data analysis. *GeoJournal, 53*(2), 149–159.

Shi, X., Duell, E., Demidenko, E., Onega, T., Wilson, B., & Hoftiezer, D. (2007). A polygon-based locally weighted average method for smoothing disease rates of small units. *Epidemiology, 18*(5), 523–528.

Slama, R., Morgenstern, V., Cyrys, J., Zutavern, A., Herbarth, O., Wichmann, H.-E., et al. (2007). Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: A study relying on a land-use regression exposure model. *Environmental Health Perspectives, 115*(9), 1283–1292.

Sohel, N., Vahter, M., Ali, M., Rahman, M., Rahman, A., Streatfield, P. K., et al. (2010). Spatial patterns of fetal loss and infant death in an arsenic-affected area in Bangladesh. *International Journal of Health Geographics, 9*, 53.

Stallones, L., Nuckols, J. R., & Berry, J. K. (1992). Surveillance around hazardous waste sites: Geographic information systems and reproductive outcomes. *Environmental Research, 59*(1), 81–92.

Talbot, T. O., Kulldorff, M., Forand, S. P., & Haley, V. B. (2000). Evaluation of spatial filters to create smoothed maps of health data, statistics in medicine. *Statistics in Medicine, 19*, 2399–2408.

Tofail, F., Vahter, M., Hamadani, J. D., Nermell, B., Huda, S. N., Yunus, M., et al. (2009). Effect of arsenic exposure during pregnancy on infant development at 7 months in rural matlab, Bangladesh. *Environmental Health Perspectives, 117*(2), 288–293.

Tu, W., Tedders, S., & Tian, J. (2011). An exploratory spatial data analysis of low birth weight prevalence in Georgia. *Applied Geography, 32*, 195–207.

Vahter, M. E., Li, L., Nermell, B., Rahman, A., El Arifeen, S., Rahman, M., et al. (2006). Arsenic exposure in pregnancy: A population-based study in Matlab, Bangladesh. *Journal of Health, Population, and Nutrition, 24*, 236–245.

von Ehrenstein, O. S., Mazumder, D. N. G., Hira-Smith, M., Ghosh, N., Yuan, Y., Windham, G., et al. (2006). Pregnancy outcomes, infant mortality, and arsenic in drinking water in West Bengal, India. *American Journal of Epidemiology, 163*, 662–669.

Waller, L. A., & Gotway, C. A. (2004). *Applied spatial statistics for public health data*. Hoboken, NJ: Wiley.

WHO. (2002). *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*. Geneva: World Health Organization.

Wilcox, A. J. (2001). On the importance—and the unimportance—of birthweight. *International Journal of Epidemiology, 30*, 1233–1241.

Xiang, H., Nuckols, J. R., & Stallones, L. (2000). A geographic information assessment of birth weight and crop production patterns around mother's residence. *Environmental Research Section A, 82*, 160–167.

Yang, C.-Y., Chang, C.-C., Tsai, S.-S., Chuang, H.-Y., Ho, C.-K., & Wu, T.-N. (2003). Arsenic in drinking water and adverse pregnancy outcome in an arseniasis-endemic area in northeastern Taiwan. *Environmental Research, 91*, 29–34.