### Title
Arsenic in private well water and birth outcomes in the United States.




### Authors
Bulka, Catherine M
Scannell Bryan, Molly
Lombard, Melissa A
et al.



### Publication Date
2022-05-01

### DOI
10.1016/j.envint.2022.107176

Peer reviewed

HHS Public Access
Author manuscript
*Environ Int*. Author manuscript; available in PMC 2023 May 01.

Published in final edited form as:
*Environ Int*. 2022 May ; 163: 107176. doi:10.1016/j.envint.2022.107176.

# Arsenic in Private Well Water and Birth Outcomes in the United States

**Catherine M. Bulka**[1], **Molly Scannell Bryan**[2], **Melissa A. Lombard**[3], **Scott M. Bartell**[4,5], **Daniel K. Jones**[6], **Paul M. Bradley**[7], **Veronica M. Vieira**[8], **Debra T. Silverman**[9], **Michael Focazio**[10], **Patricia L. Toccalino**[11], **Johnni Daniel**[12], **Lorraine C. Backer**[13], **Joseph D. Ayotte**[14], **Matthew O. Gribble**[15], **Maria Argos**[16]

[1]Department of Environmental Sciences and Engineering, University of North Carolina, 135 Dauer Drive, Chapel Hill, NC 27599, USA;

[2]Institute for Minority Health Research, University of Illinois at Chicago, 1819 W. Polk Street, Chicago, IL 60612, USA;

[3]U.S. Geological Survey, New England Water Science Center, 331 Commerce Way, Pembroke, NH 03275, USA;

[4]Department of Environmental and Occupational Health, University of California, 653 E. Peltason Drive, Irvine, CA 92697, USA;

[5]Department of Statistics, University of California, Bren Hall 2019, Irvine, CA 92697, USA;

[6]U.S. Geological Survey, Utah Water Science Center, 2329 West Orton Circle West Valley City, UT 84119, USA;

[7]U.S. Geological Survey, South Atlantic Water Science Center, 720 Gracern Rd, Columbia, SC 29210, USA;

[8]Department of Environmental and Occupational Health, University of California, 653 E. Peltason Drive, Irvine, CA 92697, USA;

[9]Occupational and Environmental Epidemiology Branch, National Cancer Institute, 9609 Medical Center Drive, Rockville, Maryland 20850, USA;

[10]U.S. Geological Survey, National Center, 12201 Sunrise Valley Dr, Reston, VA 20192, USA;

[11]U.S. Geological Survey, Northwest-Pacific Region, 2130 SW 5th Ave, Portland, OR 97201, USA;

[16]Corresponding Author: Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, 1603 West Taylor Street, Office 878A, Chicago, IL 60612, USA; argos@uic.edu.



Declaration of interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

[12]National Center for Environmental Health, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Atlanta, GA 30341, USA;

[13]National Center for Environmental Health, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Atlanta, GA 30341, USA;

[14]U.S. Geological Survey, New England Water Science Center, 331 Commerce Way, Pembroke, NH 03275, USA;

[15]Department of Epidemiology, University of Alabama at Birmingham, 217G Ryals Public Health Building, 1665 University Boulevard, Birmingham AL 35294, USA;

## Abstract

**Background:** Prenatal exposure to drinking water with arsenic concentrations >50 μg/L is associated with adverse birth outcomes, with inconclusive evidence for concentrations ≤50 μg/L. In a collaborative effort by public health experts, hydrologists, and geologists, we used published machine learning model estimates to characterize arsenic concentrations in private wells —federally unregulated for drinking water contaminants—and evaluated associations with birth outcomes throughout the conterminous U.S.

**Methods:** Using several machine learning models, including boosted regression trees (BRT) and random forest classification (RFC), developed from measured groundwater arsenic concentrations of ~20,000 private wells, we characterized the probability that arsenic concentrations occurred within specific ranges in groundwater. Probabilistic model estimates and private well usage data were linked by county to all live birth certificates from 2016 (n=3.6 million). We evaluated associations with gestational age and term birth weight using mixed-effects models, adjusted for potential confounders and incorporated random intercepts for spatial clustering.

**Results:** We generally observed inverse associations with term birth weight. For instance, when using BRT estimates, a 10-percentage point increase in the probability that private well arsenic concentrations exceeded 5 μg/L was associated with a −1.83 gram (95% CI: −3.30, −0.38) lower term birth weight after adjusting for covariates. Similarly, a 10-percentage point increase in the probability that private well arsenic concentrations exceeded 10 μg/L was associated with a −2.79 gram (95% CI: −4.99, −0.58) lower term birth weight. Associations with gestational age were null.

**Conclusion:** In this largest epidemiologic study of arsenic and birth outcomes to date, we did not observe associations of modeled arsenic estimates in private wells with gestational age and found modest inverse associations with term birth weight. Study limitations may have obscured true associations, including measurement error stemming from a lack of individual-level information on primary water sources, water arsenic concentrations, and water consumption patterns.

### Keywords

arsenic; private wells; water contamination; birth outcomes; epidemiology

## Introduction

Within the United States (U.S.), the Environmental Protection Agency (EPA) and public water purveyors work together in adherence to the Safe Drinking Water Act to make water safe for public consumption (“Safe Drinking Water Act.,” 1974). These efforts address levels of toxic chemicals, including arsenic, a metalloid with wide-ranging health effects (Agency for Toxic Substances and Disease Registry, 2007). Since 2006, regular monitoring and the use of various treatment technologies have demonstrably reduced exposure among Americans relying on public drinking water supply (Nigra et al., 2017) by maintaining arsenic concentrations below the regulatory standard of 10 μg/L (“National Primary Drinking Water Regulations; Arsenic and Clarifications to Compliance and New Source Contaminants Monitoring,” 2001). However, exposure to arsenic through drinking water has remained unchanged among private well users (Nigra et al., 2017), who comprise approximately 14% of the U.S. population (Dieter et al., 2018). Unlike public water systems, private wells are not regulated for their arsenic concentrations by the EPA nor by most states. Consequently, limited data is available on the extent of arsenic contamination and associated health risks in the U.S.

Public health experts, hydrologists, and geologists recently developed machine learning models to characterize arsenic levels in private wells throughout the conterminous U.S. (Lombard et al. 2020). The primary goal was to develop national-scale estimates of arsenic in private well water for linkage with human health data. The machine learning models offer an advantage over the traditional multivariable logistic regression model developed by Ayotte et al. (2017) by increasing sensitivity and specificity (Joseph D. Ayotte, Laura Medalie, Sharon L. Qi, Lorraine C. Backer, & Bernard T. Nolan, 2017). We used the new national well-water arsenic estimates to examine associations with gestational age and birth weight, further investigating arsenic’s role in adverse birth outcomes (Vahter, 2009). Arsenic readily crosses the placental barrier from the maternal to the fetal circulatory system (Concha, Vogler, Lezcano, Nermell, & Vahter, 1998), and accumulating evidence suggests chronic exposures might reduce fetal growth and shorten the duration of gestation (Gilbert-Diamond, Emond, Baker, Korrick, & Karagas, 2016; Howe et al., 2020; Huyck et al., 2007; Kile et al., 2016; Milton et al., 2017; Rahman et al., 2009; Xu et al., 2011; Yang et al., 2003). However, the existing epidemiologic research has been limited to small study populations, many of which were outside the U.S. and potentially subject to different exposure levels and sociocultural factors (e.g., access to adequate prenatal care, nutritional status). Given the potential public health impact of arsenic in private wells across the U.S., large, high-quality epidemiologic studies are necessary to assess the risk for adverse birth outcomes.

## Methods

### Study Population

We used the restricted-use data from the 2016 U.S. Natality File, obtained from the National Center for Health Statistics (National Center for Health Statistics 2016). Our population of interest was live, singleton births without congenital anomalies born to mothers residing in the conterminous U.S. We included births occurring from January 1 through December 31, 2016. We focused on 2016 data because this was the first year that all states implemented the 2003 revision to the U.S. Standard Certificate of Live Birth, which standardized key maternal sociodemographic items and included the obstetric estimate as a new measure for gestational age (National Center for Health Statistics, 2003). We restricted our analysis to birth records within plausible ranges of birth weight (500–5,500 grams) and gestational age (20–44 weeks). Of the 3,751,755 eligible births, we excluded mother-infant pairs with missing data on any relevant covariate (n=171,000, <5% of the full sample) to conduct a complete-case analysis. Our final analytic sample was 3,580,755 births across 3,105 counties. To focus more directly on the potential effects of chronic private well arsenic exposure on fetal growth not mediated through gestational age, we further restricted birth weight analyses to 3,305,090 term births born at 37 weeks or later (Wilcox, 2001).

Participant consent was not required for this study, as data were obtained from vital records issued by state governments to record the birth of every child within the U.S. for legal purposes. We received special approval from the National Association of Public Health Statistics and Information Systems to analyze data collected through birth certificates, including maternal residential county, which were provided in a restricted-use birth data file. The Institutional Review Board of the University of Illinois at Chicago approved this study.

### Outcome Ascertainment

Our two study outcomes of interest were gestational age and term birth weight. We investigated gestational age according to an obstetric estimate, which has greater validity than estimates based on the mother's last menstrual period (Martin, Osterman, Kirmeyer, & Gregory, 2015). Briefly, the obstetric estimate is defined as the birth attendant's best and final estimate of the infant's gestation in completed weeks (National Center for Health Statistics, 2003). We also investigated birth weight, expressed in grams, among term births to diminish the contribution of gestational age to fetal growth (Wilcox, 2001).

### Arsenic in Private Well Water

We obtained probabilistic estimates of total arsenic concentrations exceeding certain thresholds for 1 $km^2$ grids across the conterminous U.S. (Lombard et al., 2021). These estimates were produced from four distinct models: 1) a boosted regression tree (BRT) for the probability of arsenic concentrations >1 μg/L; 2) a BRT for the probability of arsenic >5 μg/L; 3) a BRT for the probability of arsenic >10 μg/L, which is the current U.S. EPA regulatory standard for public water systems ("National Primary Drinking Water Regulations; Arsenic and Clarifications to Compliance and New Source Contaminants Monitoring," 2001); and 4) a multivariate random forest classification (RFC) for the probabilities of arsenic concentrations >5 to ≤ 10 or >10 μg/L. Although the 1 μg/L cutpoint is arbitrary, 5 μg/L was selected because it is the current maximum contaminant level for public water systems in New Jersey and New Hampshire, while 10 μg/L is the current maximum contaminant level nationally ("National Primary Drinking Water Regulations; Arsenic and Clarifications to Compliance and New Source Contaminants Monitoring," 2001).

The underlying data used to develop these models included total arsenic concentrations from a total of 20,450 private wells sampled between 1970 and 2013 (J. D. Ayotte, L. Medalie, S. L. Qi, L. C. Backer, & B. T. Nolan, 2017). Of the 20,450 samples, 18,700 were obtained from the United States Geological Survey National Water Information System, 1,000 were obtained from private wells in Minnesota, and 750 were obtained from private wells in Maine (J. D. Ayotte et al., 2017). All samples were collected prior to passing through any water treatment systems. The laboratories and methods used to measure total arsenic concentrations varied, as did reporting limits. In general, earlier samples were tested with atomic absorption spectrophotometry that had reporting limits ranging from 0.9 to 1.0 μg/L, whereas more recent samples were tested using inductively coupled plasma-mass spectrometry with much lower reporting limits (<0.1 μg/L) (Garbarino, 2000). Due to the high prevalence of samples with concentrations below the respective reporting limit (9,293 samples, 45%), the machine learning models were developed to estimate the probability of exceeding a concentration threshold or occurring within a concentration range, rather than a total arsenic concentration value (Lombard et al., 2021). It should be noted that in groundwater, most arsenic is present in inorganic forms, such as trivalent arsenite ($As^3$) and pentavalent arsenate ($As^5$), rather than organic forms (Shankar, Shanker, & Shikha, 2014). A total of 249 geologic, geochemical, hydrologic, and climatic variables from various data sources were considered as candidate independent variables in the model development, as previously described (Lombard et al., 2021). Independent variables in the model varied over time; however, since a prior analysis of arsenic in repeated private well samples suggested low temporal variability (Ayotte et al., 2015), the arsenic probability estimates are assumed to be time-stable.

The final models contained between 41 and 65 independent variables, with the most influential being average annual precipitation amounts from 1981 to 2010, arsenic, selenium, and phosphorus concentrations in the C soil horizon, lateral hydrologic positions for sixth-order streams, and average annual groundwater recharge rates from 1981 to 2010 (Lombard et al., 2021). By incorporating data on these important factors, the machine learning model estimates offered a much more spatially complete representation of private well arsenic levels than observations from individual wells (Lombard et al., 2021). Model accuracies, defined as the ratio of correct model estimates to measured values, were as follows: 77.2% for the BRT of arsenic >1 μg/L; 86.2% for the BRT of arsenic >5 μg/L; 91.2% for the BRT of arsenic >10 μg/L; and 82.6% for the RFC of arsenic >5–≤ 10 or >10 μg/L (Lombard et al., 2021). Additional details on the machine learning model's performance are provided in Supplemental Table 1.

The machine learning model estimates were gridded to a spatial resolution of 1 $km^2$. However, due to confidentiality requirements, the smallest geographic subdivision within the 2016 U.S. Natality File was maternal residential county. Therefore, the modeled arsenic data were aggregated to the county-level to align with the U.S. Natality File (Figure 1). To do so, we calculated the average probability of arsenic concentrations exceeding the specified thresholds across all gridded cells within the respective county borders. We also performed sensitivity analyses to examine the impact of plausible, alternative county-level arsenic assignments, as described below (see Sensitivity Analyses section).

## Covariates

Most covariate information was obtained directly from the birth records. We considered the following individual-level maternal characteristics as potential confounders: age, race/ethnicity, marital status, educational attainment, cigarette smoking status during pregnancy, and pre-pregnancy body mass index (BMI). Maternal age at the time of delivery and pre-pregnancy BMI were examined as continuous variables (years and $kg/m^2$, respectively). We categorized race/ethnicity as non-Hispanic white, non-Hispanic Black, non-Hispanic other (including mixed race), or Hispanic. We categorized education as less than a high school diploma, high school diploma or equivalent, some college, or a college degree or greater. We dichotomized marital status as married or unmarried and cigarette smoking during pregnancy as yes or no.

County-level covariates were also considered as potential confounders. First, we considered rurality/urbanicity since private well use is more common in rural areas (Johnson, Belitz, & Lombard, 2019) where there are fewer obstetric services and, consequently, higher rates of preterm birth (Kozhimannil, Hung, Henning-Smith, Casey, & Prasad, 2018). We used the 2013 Rural-Urban Continuum Codes, developed by the U.S. Department of Agriculture, to classify maternal counties of residence as follows: metropolitan with populations of any size; non-metropolitan urbanized with populations <20,000; less urbanized with populations >2,500 to 19,999; or rural with populations <2,500 (United States Department of Agriculture Economic Research Service 2013). Next, we considered exposure to fine particulate matter with diameters ≤2.5 μm ($PM_{2.5}$) as a possible negative confounder as this type of air pollution is more prevalent in urban areas (where private well use is lower) and is an established risk factor for preterm birth and low birth weight (Shah, Balkhair, & Knowledge Synthesis Group on Determinants of Preterm, 2011). We obtained county-level annual average $PM_{2.5}$ concentrations (expressed in $\mu g/m^3$) from the Environmental Public Health Tracking Network for 2016 (Centers for Disease Control and Prevention, 2016). These estimates incorporate monitored values (from the EPA's Air Quality System) and modeled predictions (from the EPA's Downscaler model). We additionally considered the role of topography, because altitude has long been recognized as a strong determinant of low birth weight (Bailey, Donnelly, Bol, Moore, & Julian, 2019). We downloaded North American elevation data from the GTOPO30 elevation dataset produced by the U.S. Geological Survey and calculated the average elevation in meters for each county (U.S. Geological Survey, 2020).

Finally, to reduce bias owing to assigning private well arsenic estimates to mothers who may have been served by public water systems, we obtained additional county-level data on water sources. The last national-scale survey of residential water sources was conducted as part of the decadal census in 1990 (U.S. Census Bureau, 1993). However, Johnson et al. (2019) developed a novel method to estimate the population served by private wells in more recent years, including 2000 and 2010 (Johnson et al., 2019). We aggregated the estimates for data linkage purposes for 2010 (the most recent year available) from gridded cells of 1 $km^2$ to the county level (Johnson & Belitz, 2019). We then divided the estimated number of private well users by the corresponding population count from the 2010 Census to calculate the proportion of private well users within each county (U.S. Census Bureau, 2011).

## Statistical Analyses

**Regression Models of Gestational Age and Term Birth Weight—**We estimated associations of modeled private well arsenic probability estimates with gestational age and term birth weight using separate multivariable linear regression models. We included nested random intercepts and an unstructured covariance matrix for maternal county and state of residence to account for residual spatial autocorrelation in birth outcomes for mothers residing within the same geographic areas after adjusting for covariates. For example, when analyzing gestational age in relation to the county-level estimated probabilities derived from the boosted regression tree model for arsenic exceeding 1 μg/L, the model form was as follows:

$$Gestational\ Age_{ijk} = \alpha + \mu_j + \mu_k + \beta_1 * Probability\left(Arsenic > 1\frac{\mu g}{L}\right) + \varepsilon_{ijk}$$

where i refers to the individual mother-infant pair, j refers to the maternal county of residence at delivery, k refers to the maternal state of residence at delivery, α refers to the fixed global intercept which represents the grand mean of gestational age across all infants, $\mu_j$ refers to the random county-specific intercept for infant gestational age, $\mu_k$ refers to the random state-specific intercept for the infant gestational age, $ß_1$ refers to the fixed slope for the county-level probability that arsenic exceeds 1 μg/L, and $\varepsilon_{ijk}$ refers to any residual error not accounted for by model covariates and random effects.

Adjusted models further incorporated maternal age, race/ethnicity, educational attainment, marital status, smoking status during pregnancy, and pre-pregnancy BMI in addition to county rurality/urbanicity, and average annual $PM_{2.5}$ concentration as covariates. A second adjusted model was constructed, which included the proportion of private well users within the county as an additional covariate. For term birth weight, a third adjusted model was fit in which county elevation was included as an additional covariate based on the existing literature suggesting high altitude reduces fetal growth (Bailey et al., 2019).

To account for the U-shaped relationships of maternal age with prematurity and low birth weight, we modeled maternal age using restricted cubic splines (with three equally-spaced knots at the $10^{th}$, $50^{th}$, and $90^{th}$ percentiles) (Fraser, Brockert, & Ward, 1995; Hoffman et al., 2007; Lao & Ho, 1997). Similarly, non-linear associations of maternal pre-pregnancy BMI with birth outcomes were also modeled using restricted cubic splines (Kosa et al., 2011; Lewandowska, 2021). All other covariates were modeled either as continuous or categorical variables, as appropriate. We scaled estimated private well arsenic probabilities so that adjusted mean differences in gestational age and term birth weight corresponded to a 10-percentage point increase. Statistical significance was assessed using an alpha level of 0.05. We conducted all analyses in R version 4.0 (R Development Core Team) and fitted all models using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015).

**Stratified Analyses—**Due to concerns about residual confounding, we fit multivariable linear regression models stratified by geographic region and rates of private well use. In primary analyses, geographic regions were categorized according to the U.S. Geological Survey Ground Water Atlas, which is based on major aquifers (U.S. Geological Survey, 2016); in supplemental analyses, we further stratified models by maternal state of residence. For cutoff points, we stratified county-level rates of private well use as follows: 0–25%, >25–50%, >50–75%, or >75%, collapsing the latter two categories when necessary due to sparse data. In addition, we explored the potential for effect modification by infant sex as previous studies of prenatal arsenic exposure and birth outcomes have reported sex-specific relationships (Gilbert-Diamond et al., 2016; Shih, Scannell Bryan, & Argos, 2020). Stratified models included maternal age, race/ethnicity, educational attainment, marital status, smoking status during pregnancy, pre-pregnancy BMI, county-level rurality/urbanicity, and average annual $PM_{2.5}$ concentration as covariates.

**Sensitivity Analyses—**We conducted a series of sensitivity analyses. First, we fit additional binomial models with preterm birth (<37 weeks gestation), term low birth weight (<2,500 grams among infants born ≥37 weeks), small for gestational age (SGA, ≤10$^{th}$ percentile given the infant's sex and gestational age), and large for gestational age (LGA, ≥90$^{th}$ percentile given the infant's sex and gestational age) as dichotomous dependent variables to examine the clinical significance of private well arsenic exposures. Second, we fit adjusted models of gestational age and term birth weight that accounted for prenatal care utilization as an additional potential confounder; we classified prenatal care utilization as inadequate, intermediate, adequate, or adequate plus using Kotelchuck's index, which considers when prenatal care began as well as the number of prenatal care visits from initiation to delivery (Kotelchuck, 1994).

Third, we re-fit the primary regression models of gestational age and term birth weight, using the proportion of the county population relying on private well water as a sampling weight instead of a model covariate; by doing so, we down-weighted mothers who, according to their county of residence, were unlikely to use private well water. However, this sensitivity analysis changed our population of interest from all eligible births occurring in 2016 to births only among mothers who were likely to be private well users.

Fourth, because we lacked individual-level exposure data, we used a multiple imputation approach with several stages to explicitly model uncertainty in the machine learning-derived private well arsenic estimates. Accounting for errors in predictor variables is a generally appropriate and rigorous approach to epidemiological studies, and such sensitivity analysis is particularly relevant for analyses based on spatial assignment of an exposure variable encountered at the individual-level but measured in aggregate (Wright & Bateson, 2005). In the first stage of this probabilistic sensitivity analysis, we generated a county-level probability distribution for a given private well's arsenic concentration falling within each category (<5, 5–10, or > 10 μg/L). These distributions were based on the multinomial probability distributions spatially predicted for each 1 $km^2$ grid cell from the RFC models (Lombard et al., 2021) and were weighted by the private well user population (Johnson et al. 2019). Then ten values (e.g., ten hypothetical wells) were sampled for each county from that county-level probability distribution, yielding ten imputed datasets. In the next stage of the analysis, we fit regression multivariable mixed-effects linear regression models for birth outcomes, fitted to each imputed dataset. Like our primary models, we adjusted for covariates and included random intercepts for maternal residential county and state. Finally, in the third stage, we used Rubin's rules to pool model parameters into summary estimates

(Rubin, 1987). These summary estimates can be interpreted as adjusted mean differences in gestational age and term birth weight when comparing births to mothers who all used private wells containing >5 to ≤10 or >10 μg/L of arsenic, relative to mothers who all used private wells containing <5 μg/L arsenic, under the assumptions of no residual confounding and no exposure misclassification.

## Results

Our analyses included records from 3,580,755 live, singleton births during 2016 across 3,105 U.S. counties, of which 92.3% were term (Table 1). Among term births, the average birth weight was 3,383 ± 459 grams. There was an approximately even split between male and female infants. Most mothers were between the ages of 25 to 29, married, non-Hispanic white, and college-educated. Only 7.3% reported cigarette smoking while pregnant. At the time of delivery, the majority of mothers resided in metropolitan counties (86.3%) with low rates of private well use (85.0% of mothers lived in counties where private wells were used by <25% of the population)

### Private Well Arsenic and Gestational Age

Associations of private well arsenic probabilities and gestational age were largely null (Table 2). Null results were consistently observed when additionally adjusting for prenatal care utilization (data not shown), when focusing on those living in counties with high private well use by weighting (Supplemental Table 2), and when simulating private well arsenic exposures (Supplemental Table 3). Of note, models stratified by rates of private well use suggested a positive association between higher arsenic and gestational age among individuals residing in counties in which more than 75% of the population relied on private wells (Table 3); however, this level of reliance on private well water was rare, applying to only 10,348 births across 76 counties. When additionally stratifying by geographic region (Figure 3) or state (Supplemental Figure 1), some significant inverse and positive associations were observed, but most estimates were close to the null value.

### Private Well Arsenic and Preterm Birth

A total of 275,665 (7.7%) births were classified as preterm. In binomial regression models that were adjusted for individual-level and county-level covariates, and included random intercepts for maternal county and state, risk differences for private well arsenic probabilities were all zero (Supplemental Table 4).

### Private Well Arsenic and Term Birth Weight

Among the 3,305,090 term births, increasing probabilities of private well arsenic exceeding the various thresholds were generally related to lower birth weights, although some confidence intervals included the null value (Table 2); specifically, arsenic probabilities derived from boosted regression trees were significantly associated with lower term birth weights while associations with probabilities derived from the random forest classification model were imprecise. The addition of prenatal care utilization to the models did not appreciably change associations (data not shown); whereas, the addition of county-level rates of private well use and elevation to the models attenuated associations with term birth weight (see **Adjusted Models 2–3**, Table 2). Associations between increasing arsenic probabilities and term birth weight were most pronounced among male infants as compared to females (Figure 2). While weighted models suggested higher probabilities of arsenic concentrations exceeding the various thresholds were associated with lower term birth weights (Supplemental Table 2), there was no clear pattern with term birth weight in models stratified by county-level rates of private well use. In the stratified models, both negative and positive associations were observed, albeit with variable statistical significance (Table 4). There was also no apparent pattern observed when additionally stratifying by geographic region (Figure 4) or state (Supplemental Figure 2). After accounting for additional uncertainty in RFC-derived private well arsenic estimates through simulation, we again found small inverse associations that were not statistically significant (Supplemental Table 3). For instance, compared to mothers expected to have private wells containing <5 μg/L of arsenic, mothers with private wells containing >10 μg/L of arsenic gave birth to babies that weighed 2.31 (95% CI: −9.12, 4.49) grams less.

### Private Well Arsenic and Term Low Birth Weight

Of the 3,305,090 term births, 79,757 (2.4%) were classified as low birth weight (i.e., <2500 grams among infants born at 37 weeks or later). In multivariable mixed-effects binomial regression models, differences in the risk of being born at term with a low birth weight for private well arsenic probabilities derived from each machine learning model were null (Supplemental Table 4).

### Private Well Arsenic and Small and Large for Gestational Age

There were 185,078 (5.0%) infants classified as SGA and 166,414 (4.6%) infants classified as LGA. Multivariable mixed-effects binomial regression models for each outcome revealed null associations with private well arsenic probabilities (Supplemental Table 3).

## Discussion

In this large-scale study leveraging data from over 3 million births across the conterminous U.S., we evaluated the relationship between modeled arsenic probabilities in private wells and two key birth outcomes: gestational age and weight. We found no association between increased probability of elevated arsenic concentrations in private well water and gestational age at birth. We found a weak inverse relationship with birth weight among infants born at term, particularly among male infants. While these associations were consistent across different machine learning models for characterizing private well arsenic levels and were robust to adjustment for most confounding variables, they were attenuated after accounting for elevation and rates of private well use. Additional analyses, including restricting to counties with high rates of private well use, stratifying by geographic region or state, and imputing private well arsenic levels, revealed inconsistent associations between private well arsenic levels and term birth weight. Furthermore, we observed null associations in binomial models of term low birth weight, defined as a birth weight of less than 2,500 grams among infants born at or after 37 weeks gestation. Although we observed that higher private well arsenic levels were associated with lower term birth weights in some models, these associations appear to be only modest. Furthermore, the majority of mothers live in counties where private well use is rare.

Our finding of a null association of arsenic with gestational age differs from comparable epidemiologic studies conducted within the U.S. (Almberg et al., 2017; Claus Henn et al., 2016; Shi et al., 2015), but not all (Gilbert-Diamond et al., 2016; Howe et al., 2020). Several of the aforementioned studies directly measured individual-level total arsenic exposure using biomarkers (Claus Henn et al., 2016; Gilbert-Diamond et al., 2016; Howe et al., 2020), including concentrations in maternal and infant blood and urine, which integrate arsenic from all sources. By capturing arsenic exposures from contaminated drinking water and other dietary constituents (e.g., rice) (Nachman et al., 2018), biomarker-based studies likely characterize total arsenic exposure with a higher degree of precision. In general, studies that have used arsenic exposure biomarkers have observed null associations with gestational age (Gilbert-Diamond et al., 2016; Howe et al., 2020). In contrast, Almberg et al. (2017) estimated exposures using total arsenic concentrations measured exclusively in public water systems and averaged across counties, whereas Shi et al. (2015) estimated exposures using a logistic regression model of total arsenic in private well water that was aggregated to the town-level. Despite a very similar study design, Shi et al. observed positive associations of arsenic levels in private well water with preterm birth in New Hampshire. However, some key differences in study design might explain the discrepancy in findings. Foremost, we analyzed the entire conterminous U.S., where private wells are used by approximately 14% of the population (J. D. Ayotte et al., 2017). In contrast, Shi et al. (2015) focused only on the state of New Hampshire, where the rate of private well use is 40% (Shi et al., 2015). Yet, even in our stratified models, we did not find evidence of an association between elevated levels of arsenic in private wells and reduced gestational age within New Hampshire. This may be due to differences in the models used to produce probabilistic arsenic estimates. Shi et al. (2015) used estimates from a logistic regression model specifically created for the state of New Hampshire that may capture more local-scale variability in arsenic concentrations (Ayotte, Cahillane, Hayes, & Robinson, 2012). In contrast, we used estimates derived from machine learning models operating on the national-scale. Alternatively, residual confounding may explain differences; Shi et al. did not account for several maternal factors, including race/ethnicity and socioeconomic status, that we could control for based on the rich detail provided by birth certificate records.

Our finding of a modest inverse association of arsenic with birth weight in our primary analyses is consistent with the existing literature examining drinking water arsenic using spatially aggregated data (Almberg et al., 2017; Shi et al., 2015). This relationship was most pronounced among male infants, suggesting sex-specific differences in susceptibility to the gestational toxicity of arsenic exposures (Bommarito et al., 2017). However, in our secondary analyses aimed at reducing bias and accounting for uncertainty, associations of private well arsenic and term birth weight were inconsistent. Notably, the majority of births occurred to mothers residing in counties where private wells were used by less than 25% of the population. Therefore, despite a large sample size, analyses focused on births to mothers who were more likely to use private well water may have been underpowered. Future studies could incorporate natality files subsequent to 2016 to analyze this sub-group with improved statistical power. In addition, the magnitudes of associations with term birth weight were very small (<5 grams per 10-percentage point increase in the probability of exceeding respective arsenic thresholds). Arsenic may reduce fetal growth through several biological mechanisms, including oxidative stress, inflammation, and placental abnormalities (Ahmed et al., 2011). Studies relying on arsenic biomarkers have generally observed associations between exposures with lower birth weights (Claus Henn et al., 2016; Gilbert-Diamond et al., 2016; Howe et al., 2020). Despite biologic plausibility and some previous epidemiologic evidence of associations, we did not find a strong relationship between modeled private well arsenic exceedance probabilities and term birth weight.

This study has notable limitations that likely constrained our ability to detect associations between county-level private well water arsenic exceedance probabilities with adverse birth outcomes. The primary limitation of the data used for this analysis was the lack of individual-level information on the residential water source, residential histories, drinking water arsenic concentrations, and water consumption amounts and behaviors throughout pregnancy, which introduced measurement error and selection bias. We were unable to isolate mothers who definitively drank private well water during their pregnancy, and therefore, our findings may be distorted by including mothers who relied on community water systems instead. Additionally, any inaccuracies in the predicted probabilities of private well arsenic occurring in certain concentration ranges could also contribute to bias in the observed effect estimates. Among pregnant women, sociodemographic characteristics are determinants of water consumption behaviors, including intake amounts and the use of water filters (Forssen et al., 2007; Smith, Toledano, Wright, Raynor, & Nieuwenhuijsen, 2009). We may have partially reduced some exposure misclassification by including several sociodemographic variables as confounders in our birth outcome models. Relatedly, studies of residential mobility indicate few women move during pregnancy, and of those that do, most remain within the same county, suggesting only limited influence in this study (Bell & Belanger, 2012; Fell, Dodds, & King, 2004). The machine learning approach for estimating arsenic exceedance probabilities had some degree of inaccuracy. For instance, the BRT models for concentrations exceeding 5 and 10 μg/L had lower sensitivities than the BRT model for concentrations exceeding 1 μg/L, resulting in less robust prediction of areas of high arsenic levels in private well water (Lombard et al., 2021). Such errors may have been compounded by averaging the 1 $km^2$ estimates across counties and could have biased associations with birth outcomes towards the null (Richmond-Bryant & Long, 2020). Yet, we still observed an inverse association for arsenic concentrations exceeding 1 μg/L, which was derived from a very sensitive BRT model (74.1%). Additional limitations of this study include the inability to assess unmeasured confounders such as alcohol consumption, dietary intakes, or the use of certain supplements, which could be additional sources of arsenic exposure or important for arsenic detoxification (Gamble et al., 2006; Suhl et al., 2020). We were also unable to evaluate co-exposures to other contaminants (e.g., manganese, which often co-occurs with arsenic in groundwater) that may be associated with birth outcomes (Erickson et al., 2021).

There are also several strengths to this analysis. With over 3 million births, this is the largest study of arsenic exposure with birth outcomes to date. Although the machine learning model estimates of arsenic concentrations were derived using data from private wells, a recent publication suggests that arsenic levels in private wells and community water systems are positively correlated across the U.S. (Spaur et al., 2021), which supports the inclusion of births to mothers residing in counties with low rates of private well use in our analyses. Furthermore, individual-level data from birth certificates were available on birth outcomes and several covariates. Previous studies suggest that birth certificates are valid sources of information on maternal age, race/ethnicity, educational attainment, marital status, and infant gestational age and birth weight (DiGiuseppe, Aron, Ranbom, Harper, & Rosenthal, 2002; Northam & Knapp, 2006; Zollinger, Przybylski, & Gamache, 2006) thus limiting concerns for residual confounding. Finally, in addition to the large sample size, the study sample included the entire conterminous U.S., ensuring geographic representativeness.

In summary, among over 3 million births across more than 3,000 U.S. counties, we found mostly null associations between modeled probabilities of arsenic exceedance in private wells with adverse birth outcomes. An inability to ascertain exposures at the individual level likely contributed to the observed null findings. Epidemiologic and experimental data suggest arsenic exposure during pregnancy negatively affects fetal growth and development; therefore, additional research is warranted. Future birth cohort studies may be able to link to the private well arsenic estimates developed by Lombard et al. (2020) at a more granular geographic level to evaluate the impacts of exposure with less measurement error.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

This work was conducted as part of the Linking Environmental and Public Health Data to Evaluate Health Effects of Arsenic Exposure Working Group supported by the John Wesley Powell Center for Analysis and Synthesis, funded by the U.S. Geological Survey. Additional support from the U.S. Geological Survey’s National Water Quality Assessment Project and Environmental Health Programs is also appreciated. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The findings and conclusions in this report do not necessarily represent the official positions of the Centers for Disease Control and Prevention and the National Cancer Institute. This paper is dedicated to the memory of Marilyn O’Hara Ruiz.

## References

Agency for Toxic Substances and Disease Registry. (2007). Toxicological profile for Arsenic. Retrieved from Atlanta, GA.

Ahmed S, Khoda S. M. e., Rekha RS, Gardner RM, Ameer SS, Moore S, … Raqib R (2011). Arsenic-Associated Oxidative Stress, Inflammation, and Immune Disruption in Human Placenta and Cord Blood. Environmental Health Perspectives, 119(2), 258–264. doi:10.1289/ehp.1002086 [PubMed: 20940111]

Almberg KS, Turyk ME, Jones RM, Rankin K, Freels S, Graber JM, & Stayner LT (2017). Arsenic in drinking water and adverse birth outcomes in Ohio. Environ Res, 157, 52–59. doi:10.1016/j.envres.2017.05.010 [PubMed: 28521257]

Ayotte JD, Belaval M, Olson SA, Burow KR, Flanagan SM, Hinkle SR, & Lindsey BD (2015). Factors affecting temporal variability of arsenic in groundwater used for drinking water supply in the United States. Sci Total Environ, 505, 1370–1379. doi:10.1016/j.scitotenv.2014.02.057 [PubMed: 24650751]

Ayotte JD, Cahillane M, Hayes L, & Robinson KW (2012). Estimated probability of arsenic in groundwater from bedrock aquifers in New Hampshire, 2011. Retrieved from http://pubs.usgs.gov/sir/2012/5156/

Ayotte JD, Medalie L, Qi SL, Backer LC, & Nolan BT (2017). Estimating the High-Arsenic Domestic-Well Population in the Conterminous United States. Environmental Science & Technology, 51(21), 12443–12454. doi:10.1021/acs.est.7b02881 [PubMed: 29043784]

Ayotte JD, Medalie L, Qi SL, Backer LC, & Nolan BT (2017). Estimating the High-Arsenic Domestic-Well Population in the Conterminous United States. Environ Sci Technol, 51(21), 12443–12454. doi:10.1021/acs.est.7b02881 [PubMed: 29043784]

Bailey BA, Donnelly M, Bol K, Moore LG, & Julian CG (2019). High Altitude Continues to Reduce Birth Weights in Colorado. Matern Child Health J, 23(11), 1573–1580. doi:10.1007/s10995-019-02788-3 [PubMed: 31243627]

Bates D, Mächler M, Bolker B, & Walker S (2015). Fitting Linear Mixed-Effects Models Usinglme4. Journal of Statistical Software, 67(1), 1–48. doi:10.18637/jss.v067.i01

Bell ML, & Belanger K (2012). Review of research on residential mobility during pregnancy: consequences for assessment of prenatal environmental exposures. J Expo Sci Environ Epidemiol, 22(5), 429–438. doi:10.1038/jes.2012.42 [PubMed: 22617723]

Bommarito PA, Martin E, Smeester L, Palys T, Baker ER, Karagas MR, & Fry RC (2017). Fetal-sex dependent genomic responses in the circulating lymphocytes of arsenic-exposed pregnant women in New Hampshire. Reprod Toxicol, 73, 184–195. doi:10.1016/j.reprotox.2017.07.023 [PubMed: 28793237]

Centers for Disease Control and Prevention. (2016). Environmental Public Health Tracking Network. Annual PM 2.5 Level (Monitor + Modeled). Retrieved from https://ephtracking.cdc.gov/DataExplorer.. Retrieved August 13, 2020 https://ephtracking.cdc.gov/DataExplorer.

Claus Henn B, Ettinger AS, Hopkins MR, Jim R, Amarasiriwardena C, Christiani DC, … Wright RO (2016). Prenatal Arsenic Exposure and Birth Outcomes among a Population Residing near a Mining-Related Superfund Site. Environ Health Perspect, 124(8), 1308–1315. doi:10.1289/ehp.1510070 [PubMed: 26859631]

Concha G, Vogler G, Lezcano D, Nermell B, & Vahter M (1998). Exposure to inorganic arsenic metabolites during early human development. Toxicol Sci, 44(2), 185–190. doi:10.1006/toxs.1998.2486 [PubMed: 9742656]

Dieter CA, Maupin MA, Caldwell RR, Harris MA, Ivahnenko TI, Lovelace JK, … Linsey KS (2018). Estimate use of water in the United States in 2015: U.S. Geological Survey Circular Reston, VA: U.S. Geological Survey

DiGiuseppe DL, Aron DC, Ranbom L, Harper DL, & Rosenthal GE (2002). Reliability of birth certificate data: a multi-hospital comparison to medical records information. Matern Child Health J, 6(3), 169–179. doi:10.1023/a:1019726112597 [PubMed: 12236664]

Erickson ML, Elliott SM, Brown CJ, Stackelberg PE, Ransom KM, Reddy JE, & Cravotta CA (2021). Machine-Learning Predictions of High Arsenic and High Manganese at Drinking Water Depths of the Glacial Aquifer System, Northern Continental United States. Environ Sci Technol, 55(9), 5791–5805. doi:10.1021/acs.est.0c06740 [PubMed: 33822585]

Fell DB, Dodds L, & King WD (2004). Residential mobility during pregnancy. Paediatr Perinat Epidemiol, 18(6), 408–414. doi:10.1111/j.1365-3016.2004.00580.x [PubMed: 15535816]

Forssen UM, Herring AH, Savitz DA, Nieuwenhuijsen MJ, Murphy PA, Singer PC, & Wright JM (2007). Predictors of use and consumption of public drinking water among pregnant women. J Expo Sci Environ Epidemiol, 17(2), 159–169. doi:10.1038/sj.jes.7500488 [PubMed: 16670711]

Fraser AM, Brockert JE, & Ward RH (1995). Association of young maternal age with adverse reproductive outcomes. N Engl J Med, 332(17), 1113–1117. doi:10.1056/NEJM199504273321701 [PubMed: 7700283]

Gamble MV, Liu X, Ahsan H, Pilsner JR, Ilievski V, Slavkovich V, … Graziano JH (2006). Folate and arsenic metabolism: a double-blind, placebo-controlled folic acid-supplementation trial in Bangladesh. Am J Clin Nutr, 84(5), 1093–1101. doi:10.1093/ajcn/84.5.1093 [PubMed: 17093162]

Garbarino JR (2000). Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory-Determination of whole-water recoverable arsenic, boron, and vanadium using inductively coupled plasma-mass spectrometry (99–464). Retrieved from http://pubs.er.usgs.gov/publication/ofr99464

Gilbert-Diamond D, Emond JA, Baker ER, Korrick SA, & Karagas MR (2016). Relation between in Utero Arsenic Exposure and Birth Outcomes in a Cohort of Mothers and Their Newborns from New Hampshire. Environ Health Perspect, 124(8), 1299–1307. doi:10.1289/ehp.1510065 [PubMed: 26955061]

Hoffman MC, Jeffers S, Carter J, Duthely L, Cotter A, & Gonzalez-Quintero VH (2007). Pregnancy at or beyond age 40 years is associated with an increased risk of fetal death and other adverse outcomes. Am J Obstet Gynecol, 196(5), e11–13. doi:10.1016/j.ajog.2006.10.862

Howe CG, Farzan SF, Garcia E, Jursa T, Iyer R, Berhane K, … Breton CV (2020). Arsenic and birth outcomes in a predominately lower income Hispanic pregnancy cohort in Los Angeles. Environ Res, 184, 109294. doi:10.1016/j.envres.2020.109294 [PubMed: 32145549]

Huyck KL, Kile ML, Mahiuddin G, Quamruzzaman Q, Rahman M, Breton CV, … Christiani DC (2007). Maternal arsenic exposure associated with low birth weight in Bangladesh. J Occup Environ Med, 49(10), 1097–1104. doi:10.1097/JOM.0b013e3181566ba0 [PubMed: 18000415]

Johnson TD, & Belitz K (2019). Domestic well locations and populations served in the contiguous U.S.: datasets for decadal years 2000 and 2010: U.S. Geological Survey data release Retrieved from 10.5066/P9FSLU3B

Johnson TD, Belitz K, & Lombard MA (2019). Estimating domestic well locations and populations served in the contiguous U.S. for years 2000 and 2010. Sci Total Environ, 687, 1261–1273. doi:10.1016/j.scitotenv.2019.06.036 [PubMed: 31412460]

Kile ML, Cardenas A, Rodrigues E, Mazumdar M, Dobson C, Golam M, … Christiani DC (2016). Estimating Effects of Arsenic Exposure During Pregnancy on Perinatal Outcomes in a Bangladeshi Cohort. Epidemiology, 27(2), 173–181. doi:10.1097/EDE.0000000000000416 [PubMed: 26583609]

Kosa JL, Guendelman S, Pearl M, Graham S, Abrams B, & Kharrazi M (2011). The association between pre-pregnancy BMI and preterm delivery in a diverse southern California population of working women. Matern Child Health J, 15(6), 772–781. doi:10.1007/s10995-010-0633-4 [PubMed: 20602159]

Kotelchuck M (1994). The Adequacy of Prenatal Care Utilization Index: its US distribution and association with low birthweight. Am J Public Health, 84(9), 1486–1489. doi:10.2105/ajph.84.9.1486 [PubMed: 8092377]

Kozhimannil KB, Hung P, Henning-Smith C, Casey MM, & Prasad S (2018). Association Between Loss of Hospital-Based Obstetric Services and Birth Outcomes in Rural Counties in the United States. JAMA, 319(12), 1239–1247. doi:10.1001/jama.2018.1830 [PubMed: 29522161]

Lao TT, & Ho LF (1997). The obstetric implications of teenage pregnancy. Hum Reprod, 12(10), 2303–2305. doi:10.1093/humrep/12.10.2303 [PubMed: 9402300]

Lewandowska M (2021). Maternal Obesity and Risk of Low Birth Weight, Fetal Growth Restriction, and Macrosomia: Multiple Analyses. Nutrients, 13(4). doi:10.3390/nu13041213

Lombard MA, Bryan MS, Jones DK, Bulka C, Bradley PM, Backer LC, … Ayotte JD (2021). Machine Learning Models of Arsenic in Private Wells Throughout the Conterminous United States As a Tool for Exposure Assessment in Human Health Studies. Environ Sci Technol, 55(8), 5012–5023. doi:10.1021/acs.est.0c05239 [PubMed: 33729798]

Martin JA, Osterman MJ, Kirmeyer SE, & Gregory EC (2015). Measuring Gestational Age in Vital Statistics Data: Transitioning to the Obstetric Estimate. Natl Vital Stat Rep, 64(5), 1–20. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26047089

Milton AH, Hussain S, Akter S, Rahman M, Mouly TA, & Mitchell K (2017). A Review of the Effects of Chronic Arsenic Exposure on Adverse Pregnancy Outcomes. Int J Environ Res Public Health, 14(6). doi:10.3390/ijerph14060556

Nachman KE, Punshon T, Rardin L, Signes-Pastor AJ, Murray CJ, Jackson BP, … Karagas MR (2018). Opportunities and Challenges for Dietary Arsenic Intervention. Environ Health Perspect, 126(8), 84503. doi:10.1289/EHP3997 [PubMed: 30235424]

National Center for Health Statistics. (2003). Guide to Completing the Facility Worksheets for the Certificate of Live Birth and Report of Fetal Death (2003 revision). Hyattsville, MD Retrieved from http://www.cdc.gov/nchs/data/dvs/GuidetoCompleteFacilityWks.pdf.

National Center for Health Statistics (2016). User guide to the 2016 natality public use file. Retrieved from https://ftp.cdc.gov/pub/health_statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2016.pdf

National Primary Drinking Water Regulations; Arsenic and Clarifications to Compliance and New Source Contaminants Monitoring, Federal Register Volume 66, Issue 78 (April 23, 2001) Stat. (2001).

Nigra AE, Sanchez TR, Nachman KE, Harvey D, Chillrud SN, Graziano JH, & Navas-Acien A (2017). The effect of the Environmental Protection Agency maximum contaminant level on arsenic exposure in the USA from 2003 to 2014: an analysis of the National Health and Nutrition Examination Survey (NHANES). Lancet Public Health, 2(11), e513–e521. doi:10.1016/S2468-2667(17)30195-0 [PubMed: 29250608]

Northam S, & Knapp TR (2006). The reliability and validity of birth certificates. J Obstet Gynecol Neonatal Nurs, 35(1), 3–12. doi:10.1111/j.1552-6909.2006.00016.x

Rahman A, Vahter M, Smith AH, Nermell B, Yunus M, El Arifeen S, … Ekstrom EC (2009). Arsenic exposure during pregnancy and size at birth: a prospective cohort study in Bangladesh. Am J Epidemiol, 169(3), 304–312. doi:10.1093/aje/kwn332 [PubMed: 19037006]

Richmond-Bryant J, & Long TC (2020). Influence of exposure measurement errors on results from epidemiologic studies of different designs. J Expo Sci Environ Epidemiol, 30(3), 420–429. doi:10.1038/s41370-019-0164-z [PubMed: 31477780]

Rubin DB (1987). Multiple Imputation for Nonresponse in Surveys.

Safe Drinking Water Act, (1974).

Shah PS, Balkhair T, & Knowledge Synthesis Group on Determinants of Preterm, L. B. W. b. (2011). Air pollution and birth outcomes: a systematic review. Environ Int, 37(2), 498–516. doi:10.1016/j.envint.2010.10.009 [PubMed: 21112090]

Shankar S, Shanker U, & Shikha. (2014). Arsenic contamination of groundwater: a review of sources, prevalence, health risks, and strategies for mitigation. ScientificWorldJournal, 2014, 304524. doi:10.1155/2014/304524 [PubMed: 25374935]

Shi X, Ayotte JD, Onda A, Miller S, Rees J, Gilbert-Diamond D, … Moeschler J (2015). Geospatial association between adverse birth outcomes and arsenic in groundwater in New Hampshire, USA. Environ Geochem Health, 37(2), 333–351. doi:10.1007/s10653-014-9651-2 [PubMed: 25326895]

Shih YH, Scannell Bryan M, & Argos M (2020). Association between prenatal arsenic exposure, birth outcomes, and pregnancy complications: An observational study within the National Children's Study cohort. Environ Res, 183, 109182. doi:10.1016/j.envres.2020.109182 [PubMed: 32058141]

Smith RB, Toledano MB, Wright J, Raynor P, & Nieuwenhuijsen MJ (2009). Tap water use amongst pregnant women in a multi-ethnic cohort. Environ Health, 8 Suppl 1, S7. doi:10.1186/1476-069X-8-S1-S7 [PubMed: 20102592]

Spaur M, Lombard MA, Ayotte JD, Harvey DE, Bostick BC, Chillrud SN, … Nigra AE (2021). Associations between private well water and community water supply arsenic concentrations in the conterminous United States. Sci Total Environ, 787, 147555. doi:10.1016/j.scitotenv.2021.147555 [PubMed: 33991916]

Suhl J, Conway KM, Rhoads A, Langlois PH, Feldkamp ML, Michalski A, … National Birth Defects Prevention, S. (2020). Pre-pregnancy dietary arsenic consumption among women in the United States. Birth Defects Res, 112(3), 270–277. doi:10.1002/bdr2.1634 [PubMed: 31863647]

U.S. Census Bureau. (1993). 1990 Census of Housing: Detailed Housing Characteristics. Retrieved from https://www2.census.gov/library/publications/decennial/1990/ch-2/ch-2-1.pdf

U.S. Census Bureau. (2011). Summary File 1 Dataset (SF 1). Retrieved from: https://www.census.gov/data/datasets/2010/dec/summary-file-1.html

U.S. Geological Survey. (2016). Ground Water Atlas of the United States. Retrieved from https://pubs.usgs.gov/ha/ha730/gwa.html

U.S. Geological Survey. (2020). North America Elevation 1-Kilometer Resolution GRID. Retrieved from https://www.sciencebase.gov/catalog/item/4fb5495ee4b04cb937751d6d

Vahter M (2009). Effects of arsenic on maternal and fetal health. Annu Rev Nutr, 29, 381–399. doi:10.1146/annurev-nutr-080508-141102 [PubMed: 19575603]

Wilcox AJ (2001). On the importance--and the unimportance--of birthweight. Int J Epidemiol, 30(6), 1233–1241. doi:10.1093/ije/30.6.1233 [PubMed: 11821313]

Wright JM, & Bateson TF (2005). A sensitivity analysis of bias in relative risk estimates due to disinfection by-product exposure misclassification. J Expo Anal Environ Epidemiol, 15(3), 212–216. doi:10.1038/sj.jea.7500389 [PubMed: 15226753]

Xu L, Yokoyama K, Tian Y, Piao FY, Kitamura F, Kida H, & Wang P (2011). Decrease in birth weight and gestational age by arsenic among the newborn in Shanghai, China. Nihon Koshu Eisei Zasshi, 58(2), 89–95. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21473424 [PubMed: 21473424]

Yang CY, Chang CC, Tsai SS, Chuang HY, Ho CK, & Wu TN (2003). Arsenic in drinking water and adverse pregnancy outcome in an arseniasis-endemic area in northeastern Taiwan. Environ Res, 91(1), 29–34. doi:10.1016/s0013-9351(02)00015-4 [PubMed: 12550085]

Zollinger TW, Przybylski MJ, & Gamache RE (2006). Reliability of Indiana birth certificate data compared to medical records. Ann Epidemiol, 16(1), 1–10. doi:10.1016/j.annepidem.2005.03.005 [PubMed: 16039875]

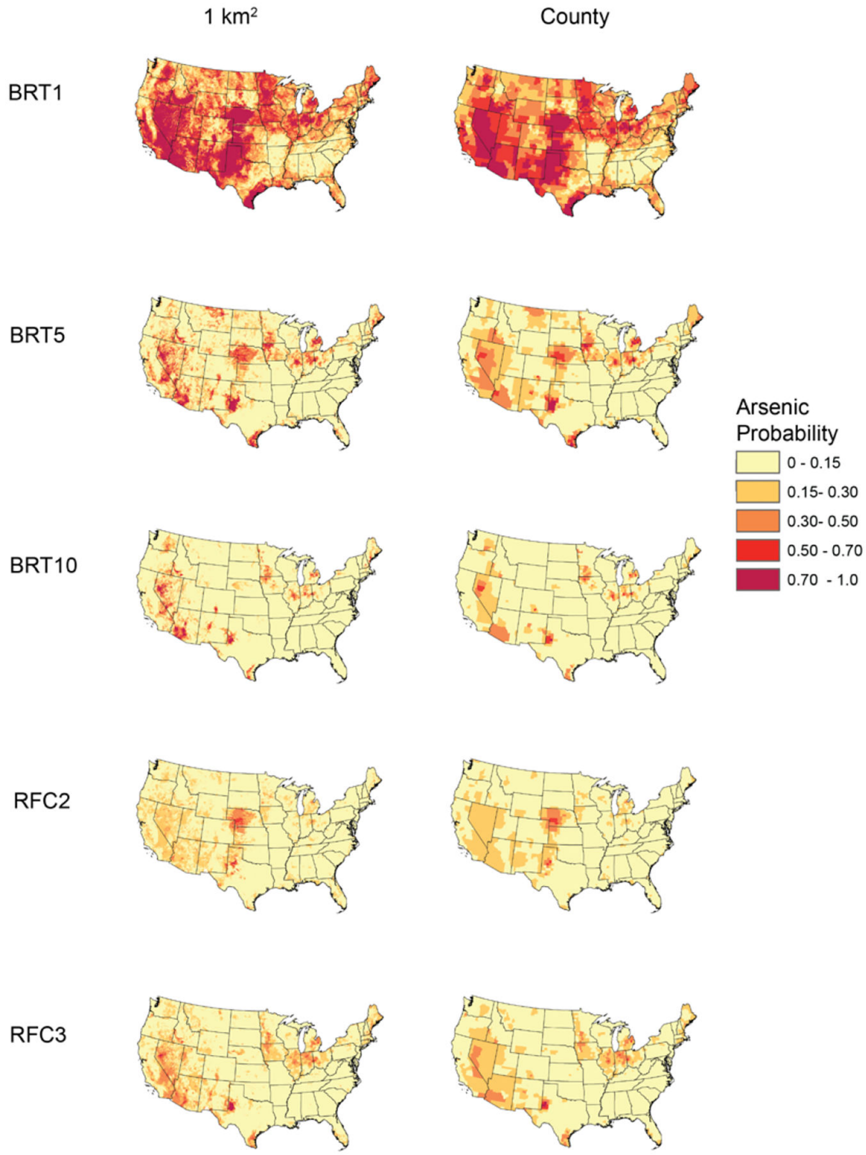


**Figure 1.**

Machine Learning Model-Predicted Arsenic Probabilities at 1 km$^2$ and Aggregated to the County-Level.

BRT1 refers to the probability that arsenic concentrations exceeded 1 μg/L as predicted by a boosted regression tree; BRT5 refers to the probability that arsenic concentrations exceeded 5 μg/L as predicted by a boosted regression tree; BRT10 refers to the probability that arsenic concentrations exceeded 10 μg/L as predicted by a boosted regression tree; RFC2 refers to the probability that arsenic concentrations fell between >5 to ≤10 μg/L whereas RFC3 refers to the probability that arsenic concentrations exceeded 10 μg/L, with both probabilities predicted by a single random forest classification model.

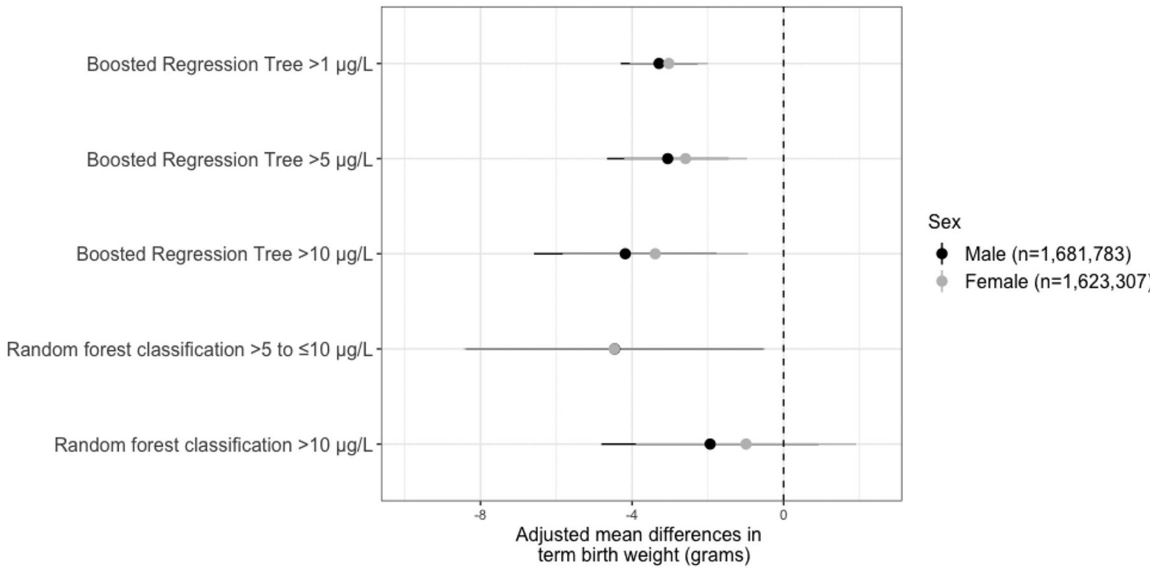


**Figure 2.**
Sex-Stratified Associations Between the Private Well Arsenic Probabilities and Term Birth Weight

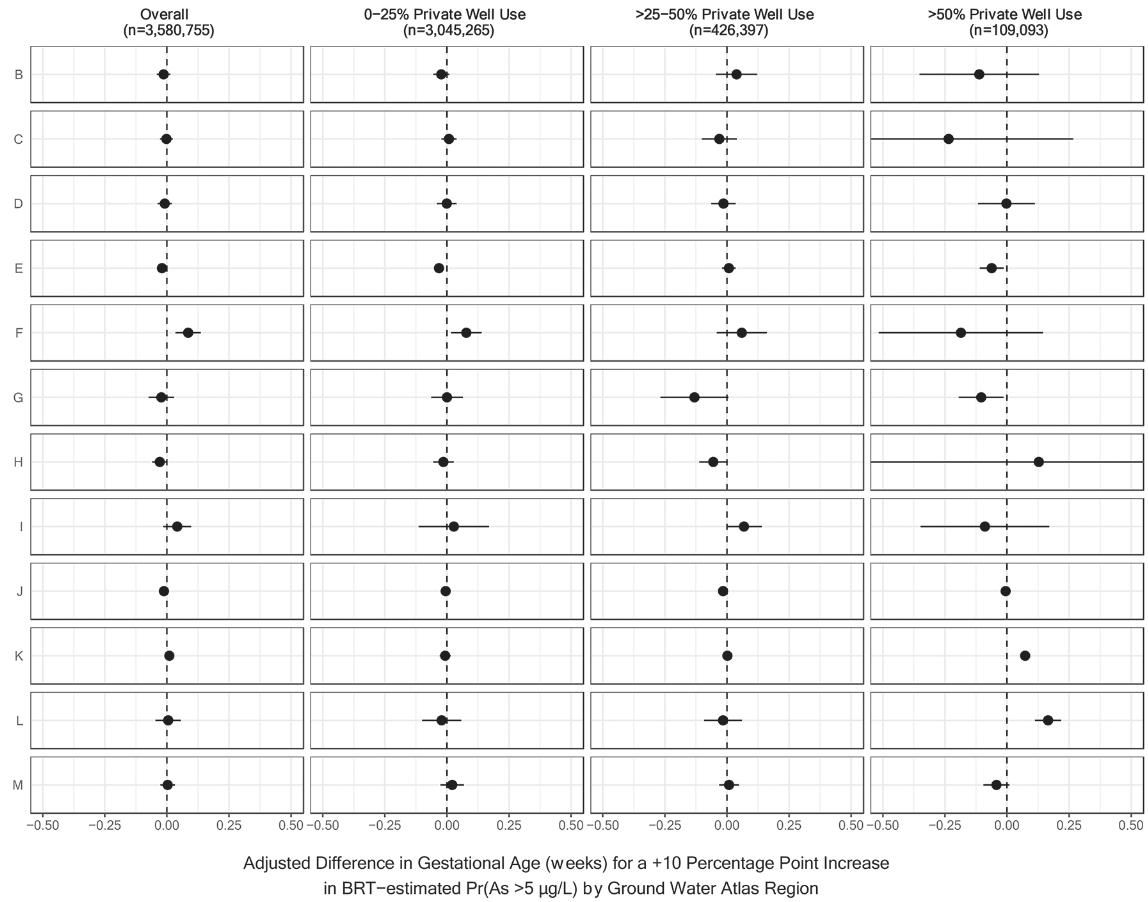


**Figure 3.**

Region-Specific Associations Between the Probability that Arsenic Concentrations Exceeded 5 μg/L and Gestational Age, Stratified by County-Level Rates of Private Well Use.

Results have been organized by U.S. Geological Survey Ground Water Atlas Regions as follows: B (California, Nevada), C (Arizona, Colorado, New Mexico, Utah), D (Kansas, Missouri, Nebraska), E (Oklahoma, Texas), F (Arkansas, Louisiana, Mississippi), G (Alabama, Florida, Georgia, South Carolina), H (Idaho, Oregon, Washington), I (Montana, North Dakota, South Dakota, Wyoming), J (Iowa, Michigan, Minnesota, Wisconsin), K (Illinois, Indiana, Kentucky, Ohio, Tennessee), L (Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, Virginia, West Virginia), and M (Connecticut, Maine, Massachusetts, New Hampshire, New York, Rhode Island, Vermont). Ground Water Atlas Regions typically share hydrogeologic and hydrologic conditions across the major aquifers in each regional area (U.S. Geological Survey, 2016). BRT-estimated Pr(As >5 μg/L) refers to the probability that private well arsenic concentrations exceeded 5 μg/L as estimated by a boosted regression tree.

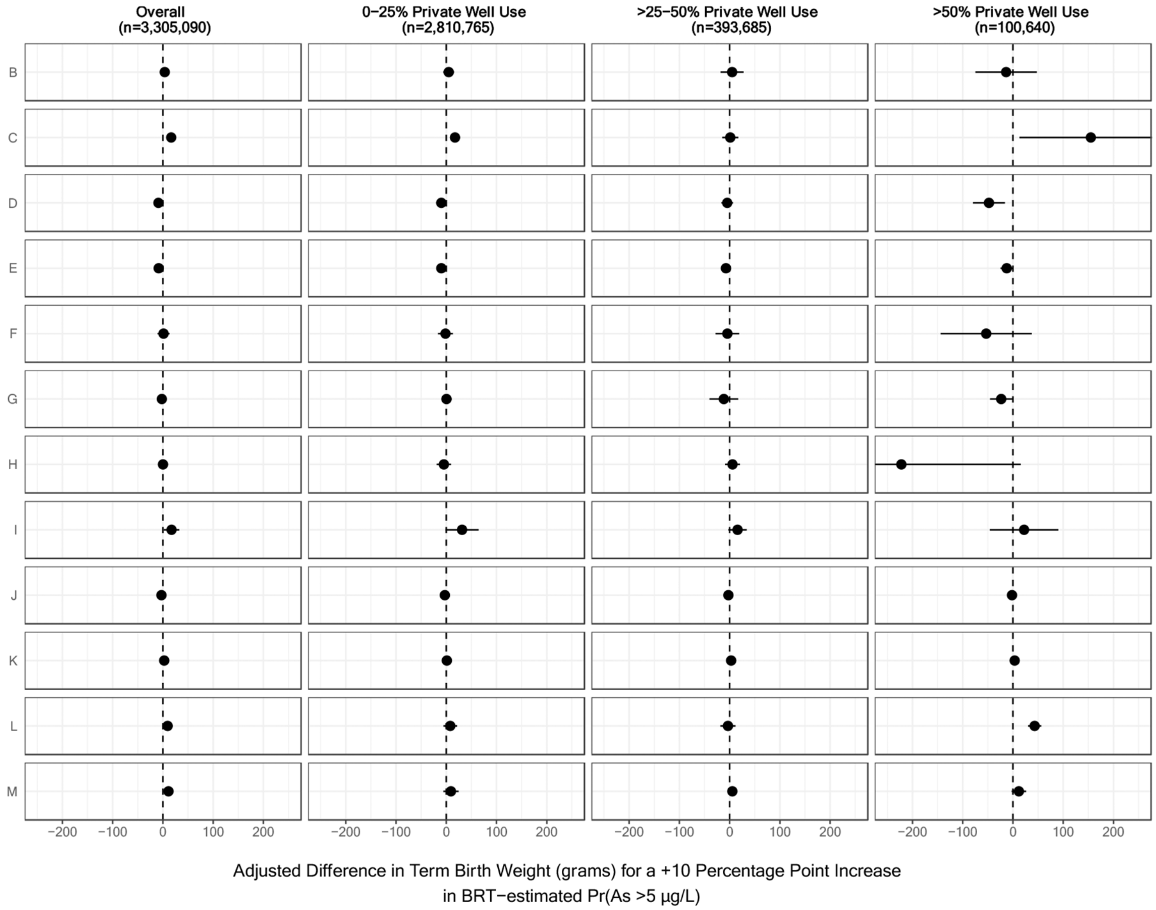


**Figure 4.**

Region-Specific Associations Between the Probability that Arsenic Concentrations Exceeded 5 μg/L and Term Birth Weight, Stratified by County-Level Rates of Private Well Use.

Results have been organized by U.S. Geological Survey Ground Water Atlas Regions as follows: B (California, Nevada), C (Arizona, Colorado, New Mexico, Utah), D (Kansas, Missouri, Nebraska), E (Oklahoma, Texas), F (Arkansas, Louisiana, Mississippi), G (Alabama, Florida, Georgia, South Carolina), H (Idaho, Oregon, Washington), I (Montana, North Dakota, South Dakota, Wyoming), J (Iowa, Michigan, Minnesota, Wisconsin), K (Illinois, Indiana, Kentucky, Ohio, Tennessee), L (Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, Virginia, West Virginia), and M (Connecticut, Maine, Massachusetts, New Hampshire, New York, Rhode Island, Vermont). Ground Water Atlas Regions typically share hydrogeologic and hydrologic conditions across the major aquifers in each regional area (U.S. Geological Survey, 2016). BRT-estimated Pr(As >5 μg/L) refers to the probability that private well arsenic concentrations exceeded 5 μg/L as estimated by a boosted regression tree.

**Table 1.**

Demographics of 3,580,755 singleton births and distribution of gestational age and term birth weight across 3,105 U.S. counties during 2016.

| Characteristic | N (%) | Gestational Age (weeks) Mean ± S.D. | Term Birth Weight (grams)[a] Mean ± S.D. |
|---|---|---|---|
| Preterm Birth | | | |
| Yes | 275,665 (7.7) | 34.1 ± 2.8 | -- |
| No | 3,305,090 (92.3) | 39.0 ± 1.1 | 3,383 ± 459 |
| Infant Sex | | | |
| Female | 1,750,363 (48.9) | 38.7 ± 1.8 | 3,318 ± 446 |
| Male | 1,830,392 (51.1) | 38.6 ± 1.9 | 3,445 ± 462 |
| Maternal Age (years) | | | |
| <25 | 936220 (26.1) | 38.6 ± 1.9 | 3,311 ± 446 |
| 25–29 | 1,050,606 (29.3) | 38.7 ± 1.8 | 3,387 ± 455 |
| 30–34 | 1,003,472 (28.0) | 38.7 ± 1.8 | 3,423 ± 460 |
| ≥35 | 590,457 (16.5) | 38.5 ± 1.9 | 3,419 ± 471 |
| Marital Status | | | |
| Unmarried | 1,424,914 (39.8) | 38.5 ± 2.0 | 3,316 ± 458 |
| Married | 2,156,741 (60.2) | 38.7 ± 1.7 | 3,426 ± 455 |
| Maternal Race/Ethnicity | | | |
| Non-Hispanic White | 1,899,029 (53.0) | 38.8 ± 1.7 | 3,441 ± 458 |
| Non-Hispanic Black | 503,702 (14.1) | 38.3 ± 2.2 | 3,236 ± 452 |
| Non-Hispanic Other | 332,343 (9.3) | 38.6 ± 1.7 | 3,303 ± 444 |
| Hispanic | 845,681 (23.6) | 38.6 ± 1.8 | 3,368 ± 445 |
| Maternal Education | | | |
| Less than high school or equivalent | 491,945 (13.7) | 38.5 ± 1.9 | 3,317 ± 460 |
| High school diploma or equivalent | 903,808 (25.2) | 38.6 ± 1.9 | 3,340 ± 459 |
| Some college | 747,156 (20.9) | 38.6 ± 1.8 | 3,383 ± 461 |
| College degree or greater | 1,437,846 (40.2) | 38.8 ± 1.7 | 3,431 ± 452 |
| Maternal Smoking During Pregnancy | | | |
| Yes | 260,479 (7.3) | 38.4 ± 2.1 | 3,219 ± 465 |

| Characteristic | N (%) | Gestational Age (weeks) Mean ± S.D. | Term Birth Weight (grams)[a] Mean ± S.D. |
|---|---|---|---|
| No | 3,320,276 (92.7) | 38.7 ± 1.8 | 3,395 ± 456 |
| Pre-pregnancy Body Mass Index (BMI) | | | |
| Underweight (<18.5 kg/m$^2$) | 126,353 (3.5) | 38.5 ± 1.9 | 3,180 ± 428 |
| Healthy weight (≥18.5–24.9 kg/m$^2$) | 1,590,485 (44.4) | 38.7 ± 1.8 | 3,343 ± 441 |
| Overweight (≥25.0–29.9 kg/m$^2$) | 932,413 (26.0) | 38.7 ± 1.8 | 3,411 ± 456 |
| Obese (≥30.0 kg/m$^2$) | 931,504 (26.0) | 38.5 ± 2.0 | 3,452 ± 481 |
| Rurality/Urbanicity | | | |
| Metropolitan | 3,089,816 (86.3) | 38.7 ± 1.8 | 3,382 ± 458 |
| Non-metropolitan urbanized | 199,487 (5.6) | 38.6 ± 1.8 | 3,388 ± 465 |
| Less urbanized | 244,500 (6.8) | 38.6 ± 1.8 | 3,384 ± 465 |
| Thinly populated | 46,952 (1.3) | 38.6 ± 1.8 | 3,397 ± 465 |
| County-Level Private Well Use | | | |
| 0–25% | 3,045,265 (85.0) | 38.6 ± 1.8 | 3,378 ± 457 |
| >25–50% | 426,397 (11.9) | 38.7 ± 1.8 | 3,409 ± 466 |
| >50–75% | 98,745 (2.8) | 38.7 ± 1.8 | 3,410±472 |
| >75% | 10,348 (0.3) | 38.7 ± 1.9 | 3,413 ± 470 |

[a] Birth weight means and standard deviations calculated only among the 3,305,090 infants born at term

## Table 2.

Mean differences in gestational age and term birth weight for a +10-percentage point increase in the probability that private well arsenic concentrations exceed the respective threshold

| | Gestational Age (weeks) | | | Term Birth Weight (grams) | | | |
|---|---|---|---|---|---|---|---|
| **No. of Births** | 3,580,755 | | | 3,305,090 | | | |
| **No. of Counties** | 3,105 | | | 3,105 | | | |
| **Private well arsenic level** | **Unadjusted[a] β (95% CI)** | **Adjusted[b] β (95% CI)** | **Adjusted 2[c] β (95% CI)** | **Unadjusted[a] β (95% CI)** | **Adjusted[b] β (95% CI)** | **Adjusted 2[c] β (95% CI)** | **Adjusted 3[d] β (95% CI)** |
| **Boosted regression tree: 1 μg/L** | | | | | | | |
| ≤1 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >1 μg/L | −0.01 (−0.01, −0.01)* | −0.01 (−0.01, −0.01)* | −0.01 (−0.01, −0.01)* | −3.13 (−4.30, −1.95)* | −3.12 (−4.05, −2.18)* | −3.10 (−4.04, −2.17)* | −2.84 (−3.74, −1.94)* |
| **Boosted regression tree: 5 μg/L** | | | | | | | |
| ≤5 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >5 μg/L | −0.01 (−0.01, 0.00) | −0.01 (−0.01, 0.00) | −0.01 (−0.01, 0.00) | −2.24 (−4.10, −0.38)* | −2.73 (−4.19, −1.26)* | −2.75 (−4.21, −1.29)* | −2.20 (−3.61, −0.78)* |
| **Boosted regression tree: 10 μg/L** | | | | | | | |
| ≤10 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >10 μg/L | −0.01 (−0.02, 0.00) | −0.01 (−0.02, 0.00) | −0.01 (−0.02, 0.00) | −3.08 (−5.91, −0.26)* | −3.85 (−6.07, −1.64)* | −3.89 (−6.10, −1.68)* | −2.99 (−5.13, −0.86)* |
| **Random forest classification: 5 and 10 μg/L** | | | | | | | |
| ≤5 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >5 to ≤10 μg/L | 0.00 (−0.01, 0.02) | 0.00 (−0.02, 0.01) | 0.00 (−0.02, 0.01) | −3.62 (−7.97, 0.73) | −4.39 (−7.91, −0.86)* | −4.37 (−7.89, −0.85)* | −1.62 (−5.06, 1.82) |
| >10 μg/L | 0.00 (−0.02, 0.01) | 0.00 (−0.01, 0.01) | 0.00 (−0.01, 0.01) | −1.49 (−4.79, 1.82) | −1.54 (−4.16, 1.06) | −1.56 (−4.16, 1.05) | −1.95 (−4.47, 0.40) |

Note: CI, confidence interval. Asterisks denote statistical significance.

[a]Unadjusted model includes the percentage of private wells in the county predicted to have arsenic concentrations within the respective category with nested random intercepts for maternal county and state of residence.

[b]Adjusted model additionally includes maternal age (modeled as a restricted cubic spline), race/ethnicity (non-Hispanic white, non-Hispanic black, non-Hispanic other, or Hispanic), marital status (unmarried or married), educational attainment (less than high school, high school diploma or equivalent, some college, or college degree or greater), smoking during pregnancy (yes or no), pre-pregnancy BMI (modeled as a restricted cubic spline), rurality/urbanicity (metropolitan, non-metropolitan, less urbanized, or thinly populated), and the average annual concentration of particulate matter with aerodynamic diameter less than 2.5 μm in the county (μg/m3).

[c]Adjusted model 2 additionally includes the proportion of private well users in each county.

[d]Adjusted model 3 additionally includes the average elevation of the county (m).

**Table 3.**

Adjusted mean differences in gestational age for a +10-percentage point increase in the probability that private well arsenic concentrations exceed the respective threshold, overall and stratified by county-level rates of private well use.

| | Gestational Age (weeks) β (95% CI) | | | | |
|---|---|---|---|---|---|
| **Private well arsenic level** | **Overall** | **0–25% private well use** | **>25–50% private well use** | **>50–75% private well use** | **>75% private well use** |
| **No. of Births** | 3,580,755 | 3,045,265 | 426,397 | 98,745 | 10,348 |
| **No. of Counties** | 3,105 | 1,522 | 1,091 | 415 | 76 |
| **Boosted regression tree: 1 μg/L** | | | | | |
| ≤1 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >1 μg/L | −0.01 (−0.01, −0.01)* | −0.01 (−0.01, −0.01)* | −0.01 (−0.02, −0.01)* | 0.00 (−0.02, 0.01) | 0.04 (−0.01, 0.09) |
| **Boosted regression tree: 5 μg/L** | | | | | |
| ≤5 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >5 μg/L | −0.01 (−0.01, 0.00) | −0.01 (−0.02, −0.01)* | −0.01 (−0.02, 0.01) | 0.00 (−0.02, 0.02) | 0.12 (0.02, 0.21)* |
| **Boosted regression tree: 10 μg/L** | | | | | |
| ≤10 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >10 μg/L | −0.01 (−0.02, 0.00) | −0.02 (−0.03, −0.01)* | 0.00 (−0.02, 0.01) | 0.01 (−0.02, 0.04) | 0.33 (0.15, 0.52)* |
| **Random forest classification: 5 and 10 μg/L** | | | | | |
| ≤5 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >5 to ≤10 μg/L | 0.00 (−0.02, 0.01) | 0.00 (−0.02, 0.02) | −0.02 (−0.04, 0.01) | 0.00 (−0.06, 0.05) | −0.09 (−0.25, 0.07) |
| >10 μg/L | 0.00 (−0.01, 0.01) | −0.01 (−0.03, 0.00) | 0.00 (−0.01, 0.02) | 0.01 (−0.02, 0.05) | 0.29 (0.15, 0.43)* |

Note: CI, confidence interval. Asterisks denote statistical significance. Estimates are adjusted for maternal age, race/ethnicity, marital status, educational attainment, smoking during pregnancy, pre-pregnancy BMI, rurality/urbanicity, and the average annual concentration of particulate matter with aerodynamic diameter less than 2.5 μm in the county with nested random intercepts for maternal county and state of residence.

**Table 4.**

Adjusted mean differences in term birth weight for a +10-percentage point increase in the probability that private well arsenic concentrations exceed the respective threshold, overall and stratified by county-level rates of private well use.

| | Term Birth Weight (grams) β (95% CI) | | | | |
|---|---|---|---|---|---|
| | **Overall** | **0–25% private well use** | **>25–50% private well use** | **>50–75% private well use** | **>75% private well use** |
| **No. of Births** | 3,305,090 | 2,810,765 | 393,685 | 91,114 | 9,526 |
| **No. of Counties** | 3,105 | 1,522 | 1,091 | 415 | 76 |
| **Boosted regression tree: 1 μg/L** | | | | | |
| ≤1 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >1 μg/L | −3.12 (−4.05, −2.18)* | 4.19 (−5.33, −3.05)* | −1.78 (−3.50, −0.06)* | −3.63 (−6.99, −0.27)* | 1.49 (−7.65, 10.60) |
| **Boosted regression tree: 5 μg/L** | | | | | |
| ≤5 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >5 μg/L | −2.73 (−4.19, −1.26)* | −4.86 (−6.81, −2.91)* | −0.61 (−3.06, 1.85) | −2.51 (−6.69, 1.68) | 3.52 (−16.30, 23.30) |
| **Boosted regression tree: 10 μg/L** | | | | | |
| ≤10 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >10 μg/L | −3.85 (−6.07, −1.64)* | −6.86 (−10.10, −3.65)* | −1.45 (−4.80, 1.91) | −2.87 (−9.01, 3.28) | 20.10 (−19.90, 60.00) |
| **Random forest classification: 5 and 10 μg/L** | | | | | |
| ≤5 μg/L | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) | 0.00 (reference) |
| >5 to ≤10 μg/L | −4.39 (−7.91, −0.86)* | −6.91 (−11.40, −2.46)* | −1.67 (−7.57, 4.24) | 1.54 (−11.70, 14.70) | −16.10 (−51.30, 19.20) |
| >10 μg/L | −1.54 (−4.16, 1.06) | −3.38 (−7.00, 0.25) | −1.33 (−5.55, 2.89) | −2.58 (−9.85, 4.69) | 32.00 (1.30, 62.80)* |

Note: CI, confidence interval. Asterisks denote statistical significance. Estimates are adjusted for maternal age, race/ethnicity, marital status, educational attainment, smoking during pregnancy, pre-pregnancy BMI, rurality/urbanicity, and the average annual concentration of particulate matter with aerodynamic diameter less than 2.5 μm in the county with nested random intercepts for maternal county and state of residence.