



# HHS Public Access

Author manuscript

*Sci Total Environ.* Author manuscript; available in PMC 2023 March 15.

Published in final edited form as:

*Sci Total Environ.* 2022 March 15; 812: 151479. doi:10.1016/j.scitotenv.2021.151479.

## Analysis of the novel NCWELL database highlights two decades of co-occurrence of toxic metals in North Carolina private well water: public health and environmental justice implications

Lauren A. Eaves<sup>1,2</sup>, Alexander P. Keil<sup>3</sup>, Julia E. Rager<sup>1,2,4</sup>, Andrew George<sup>2,5</sup>, Rebecca C. Fry<sup>1,2,4,6,†</sup>

<sup>1</sup> Department of Environmental Sciences & Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

<sup>2</sup> Institute for Environmental Health Solutions, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

<sup>3</sup> Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

<sup>4</sup> Curriculum in Toxicology and Environmental Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

<sup>5</sup> Institute for the Environment, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

<sup>6</sup> Department of Pediatrics, School of Medicine, University of North Carolina, Chapel Hill, North Carolina, USA

### Abstract

Private well users are particularly vulnerable to metal exposure as they are not protected by the Safe Drinking Water Act. In North Carolina (NC), approximately 2.4 million individuals rely on private well water. In the present study, we constructed the NCWELL database: a comprehensive database of 117,960 geocoded well water tests over twenty-years in NC inclusive of 28 metals/metalloids. The NCWELL database was analyzed to identify areas of concern for single and co-occurring toxic metal contamination of private wells in NC. County-level population-at-risk rankings were calculated by combining toxic metal levels and the proportion of residents relying

---

†Corresponding Author: rfry@unc.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.

CRediT author statement

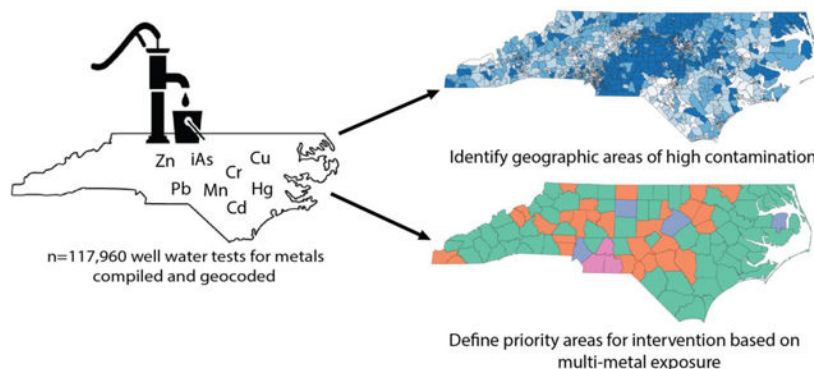
**Lauren A. Eaves:** Conceptualization, Methodology, Software, Validation, Formal analysis, Writing-Original draft preparation.

**Alexander P. Keil:** Conceptualization, Software, Data Curation, Writing-Review & Editing, Funding acquisition; **Julia E Rager:** Writing-Review & Editing; **Andrew George:** Writing-Review & Editing; **Rebecca C. Fry:** Conceptualization, Methodology, Writing-Review & Editing, Project Administration, Funding acquisition

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

on well water. Additionally, k-means analysis was used to identify counties with critical co-occurrence of toxic metals. These results highlight that inorganic arsenic (iAs) and lead (Pb) were detected above the EPA standards of 10 and 15 ppb in over 2,500 and over 3,000 tests, respectively. Shockingly, iAs was observed at levels up to 806 ppb and Pb at levels up to 105,440 ppb. Manganese (Mn) was detected above the EPA lifetime Health Advisory Limit in 4.9% and above the secondary Maximum Contaminant Level in 24.3% of all well water tests in NC, with a maximum concentration of 46,300 ppb reported. Mixtures-based analysis identified four distinct clusters of counties, one demonstrating high iAs and Mn and another with high Pb. Over the twenty-year period, metal levels remained high, indicative of sustained contamination in areas of concern. This study provides a novel database for researchers and concerned citizens in NC, demonstrates a methodology for identifying priority geographic regions for single and multiple contaminants, and has environmental justice implications in NC where metal exposure via private well water remains a serious public health concern.

## Graphical Abstract



## Keywords

Arsenic; manganese; lead; private wells; k-means clustering; geocoding

## 1. Introduction

Private wells serve as the primary drinking water source for approximately 42 million people in the United States (US) (Dieter et al. 2018). The Environmental Protection Agency (EPA) regulates public drinking water systems under the Safe Drinking Water Act which enforces drinking water standards for a variety of contaminants, including metals, such as lead (Pb) and manganese (Mn), and metalloids, such as inorganic arsenic (iAs) (Sanders et al. 2012). However, private wells are not federally regulated and therefore the water quality of over 40 million Americans is not monitored on a routine basis, leaving private well users vulnerable to microbial and chemical contamination. Numerous contaminants of concern have been detected in groundwater sources throughout the US, including Pb, iAs (technically a metalloid, referred to herein as a metal), mercury (Hg) cadmium (Cd), as well as essential metals with known adverse effects at high/chronic exposures, such as Mn (Peters

et al. 1999; Kim et al. 2002; Ayotte et al. 2003; Shaw et al. 2005; Yang et al. 2009; Sanders et al. 2012).

Prior studies suggest that well water users have a higher body burden of toxic metals, including iAs and Pb, than individuals on public water systems (Nigra et al. 2017; Gibson et al. 2020). Toxic metal exposure is of critical concern to public health, underscored by the fact that iAs, Pb, Hg and Cd are ranked as the first, second, third and seventh chemicals, respectively, on the Agency for Toxic Substances and Disease Registry's (ATSDR) substance priority list for 2019 (ATSDR 2019). Exposure to iAs, Mn or Pb is associated with a wide array of disease endpoints, including cancer. Ingestion of iAs in drinking water is associated with heart disease, neurological abnormalities, adverse reproductive outcomes (including reduced birth weight and infant mortality) as well as skin, lung, kidney, and bladder cancer (Naujokas et al. 2013). Mn, while essential in low doses, has been linked to infant mortality and birth defects as well as neurological impairment, especially when exposed early in life (Spangler and Spangler 2009; Sanders et al. 2014; O'Neal and Zheng 2015). Lastly, it is well established that there is no safe dose for Pb, with neurotoxic effects, especially for children, observed well below current guidelines (Needleman and Bellinger 1991; Levallois et al. 2018). There is also a growing body of evidence demonstrating that Pb exposure increases the risk of adverse birth outcomes, including preterm birth and preeclampsia (Andrews et al. 1994; Ferguson and Chin 2017; Poropat et al. 2018).

NC is a critical state for the evaluation of private well water contamination as it has the fifth highest population in the U.S. (2.4 million people or 24% of the overall state population) using private well water (Dieter et al. 2018). As of 2008, NC requires that all new wells be tested for bacterial and chemical contaminants within 30 days of being built, and after the initial test it is recommended that well owners test their water for bacterial contamination every year and for heavy metals, nitrates and nitrates every two years (NC Department of Public Health, Epidemiology section). While the Safe Drinking Water Act is actively enforced through public water systems annual reporting of regulated chemicals and penalties for exceedances of standards, the NC Groundwater Standards do not actively monitor the safety of private well water. Instead, the responsibility for maintaining water safety – including monitoring and remediation – falls onto the individual homeowner. By law, other than the initial test after a new well is constructed, there are no mandated testing requirements for private wells in NC.

Importantly, metal exposure mitigation technologies for private well water users in NC have proven effective, with a tabletop water filter reducing concentrations of Pb and iAs by 99% (Tomlinson et al. 2019). However, there are substantial socioeconomic, situational, and psychological barriers to both testing and treatment. At an individual well user level, misconceptions that contaminants can be detected by sensory perception, costs of testing and treatment, and lack of awareness regarding how to get a water test all present barriers to achieve safe well water (Zheng and Flanagan 2017; Stillo et al. 2019). Beyond the individual level, local health department private well user support programs are generally under resourced and exhibit substantial variability in capacity across the state (Gibson and Pieper 2017; Wait et al. 2020).

In our prior study of private wells in NC, we geocoded and mapped iAs levels in over 63,000 wells tested from 1998 to 2010 and highlighted that approximately 700 well water tests exceeded the EPA's maximum contaminant level (MCL) of 10 parts per billion (ppb) for iAs, with the highest concentration measuring over 80 times the MCL (Sanders et al. 2012). In addition, Mn was found to co-occur with iAs and was present over the EPA secondary MCL (SMCL) in 20.5% of included well water tests (Sanders et al. 2014). Additionally, Pb exceeded the EPA MCL in 3.1% (Sanders et al. 2014). Moreover, other studies and reports have found elevated levels of vanadium (V) and hexavalent chromium (CrVI) in groundwater throughout NC (Tomlinson et al. 2019; Coyte et al. 2020). The occurrence of contaminants in well water in NC, and elsewhere, can be derived from multiple possible sources including industry-derived contamination (e.g. leaking coal ash pits), corrosion in water pipes or naturally occurring concentrations related to the composition of the aquifer rocks and soil (Tomlinson et al. 2019).

The common occurrence of toxic metals, their toxicity, as well as fragmented private well user support programs, makes metal exposure via private well water an area of serious public health concern across the US, and in NC specifically. Despite this, systematic evaluations of private well water exposure in the US are sparse and an updated and comprehensive evaluation of a suite of metals in private wells across NC has not been conducted, to our knowledge, since previous work by our team (Sanders et al. 2012, 2014). Therefore, we set out to (1) provide a comprehensive database of 28 metals/metalloids tested in private well water spanning the last twenty years in NC that can be utilized by public health researchers and professionals; (2) to identify counties of concern for high priority contaminants, namely iAs, Pb and Mn (Sanders et al. 2012, 2014; Tomlinson et al. 2019); (3) to evaluate co-occurrence of metals through correlation and, for toxic metals, k-means clustering; and lastly (4) to assess whether levels of metals reported in private well water vary over time. To achieve these aims, we utilized approximately 118,000 private well water measurements collected by the North Carolina Department of Health and Human Services (NCDHHS) dating from October 19, 1998 to May 20, 2019. Our results indicate high levels of metal contamination in private wells and emphasize the need for universal screening of private wells. These results also provide information on regions in NC where critical action is required to provide safe drinking water to well water users.

## 2. Materials and Methods

### 2.1 Private well water test data collection

Private well owners can request a well water quality test through their local health department. The median cost for a well water test of inorganic analytes is \$90; however, the cost of the testing varies significantly by local health department from \$25 to over \$200 (NC Department of Public Health, Epidemiology section; Wait et al. 2020). Local health department officials collect water samples from indoor, outdoor, or well tap after letting water run for 5–10 minutes and then ship the samples to the NCDHHS Division of Public Health State Laboratory of Public Health where they are analyzed within 48 hours. Detailed analytical methods used to characterize the contaminants in well water samples are provided in Supplementary Information (Section 2.1). Test findings are summarized in reports sent

back to the well owner. The reports contain: the concentration in mg/L (ppm), or for CrVI, in µg/L (ppb), or an indication that the concentration was below a limit of reporting (LOR). The LOR is the minimum value below which no quantitative concentration is reported. The individual reports sent to well owners are publicly available on the NCDHHS website (NC Division of Public Health).

## 2.2 NCWELL database construction

One of the goals of this study was to develop the NCWELL database combining the dataset published in Sanders et al. 2012 (representative of well water tests from approximately 1998–2010, thus termed “pre-2010” data) with a dataset representative of well water tests from the last ten years (2010–2019, thus termed “post-2010” data) (Figure S1) (Sanders et al. 2012). The pre-2010 data only includes iAs, Mn, Pb and Cd, the post-2010 data includes a wider suite of metals/metalloids.

For the post-2010 data, publicly available well water test reports were downloaded as PDF forms comprising laboratory measurements on 05/23/2019 from the NCDHHS State Laboratory of Public Health, inorganic chemistry section website. The files were filtered for only those that represented well water measurements and excluded other analyses (e.g., testing municipal water for fluoride levels or testing dust samples for lead) based on key word abstraction. In total, n=78,163 distinct reports were identified, from which tabular measurement data and meta data related to the residence were abstracted from the PDF forms using the R package “pdftools,” which was facilitated by consistent PDF formatting across all study years and areas (Ooms 2021). Abstracted fields included unique state-supplied identification numbers, residential address (street, city, and ZIP code) of well location, sampling, and analysis dates, and reported concentrations of metal/metalloids measured (or that the metal/metalloid was measured below the LOR). Mecklenburg county addresses were not included in data from the state but were separately abstracted from Mecklenburg county Well Information System website and linked via a separate identification number used for Mecklenburg county records (Mecklenburg County Health Department). Data were first cleaned of parsing errors and missing data, resulting in removal of n=1,633 tests. Detailed justifications for removal are provided in Supplementary Information (Section 2.2).

For the pre-2010 data, the database was obtained through internal laboratory records and was inclusive of n=46,286 tests. Of these, n=44,767 had available street addresses that could be re-geocoded with an updated address locator (see further details below). The remaining n=1,519 had associated latitude and longitude coordinates that were originally provided by the NCDHHS.

## 2.3 Geocoding

All tests in the post-2010 data were geocoded. Additionally, tests from the pre-2010 data with available street addresses were geocoded to assess repeatability of geocoding and to apply a standardized, updated address locator. Well water tests were geocoded utilizing ArcGIS Desktop ESRI version 10.7 and ESRI’s Online World Geocoding Service address locator. In the Sanders et al. analysis a four-class geocoding scheme was created; here, we

utilized Class II geocoding, according to their definition (Sanders et al. 2012). Specifically, the residential street address reported in the well water test report, comprising street address and where available, city and zip code were used as input variables into the “US Address-Single House” style address locator to determine the most likely location match. The latitude and longitude coordinates (in decimal degrees format) were utilized in a spatial join to the census tract and county boundaries, based on the 2010 census reference data, to generate state, county, and census tract FIPS codes for each well water test. Following geocoding, addresses matched outside of NC were removed and the post-2010 and pre-2010 data were combined at this stage (n=122,130) and subsequent QA/QC of the geocoding was conducted. Details regarding the removal of addresses outside of NC are provided in Supplementary Information (Section 2.3).

Of the n=122,130, over 70% of addresses were matched to the exact point address and 15% matched to the street address. Several variables output from ArcGIS can be utilized to evaluate the precision of the geocoding match. Here, we utilized the Address type (“Addr\_type”) which details whether the address was matched to an exact point address, a street, a zip code, a point of interest, or another geographic unit (“Locality”). If the match is to a street, a zip code, a point of interest, or another geographic unit, the coordinates of the centroid are provided. The confidence scores (“Score”), which provides a numeric score from 0–100 of the confidence in the match, were also examined. Any test that was matched to a Locality was removed if the confidence score was <80 or the locality unit was “State or Region” as the geocoding match was unlikely to be sufficiently accurate. This filter resulted in the removal of n=1,286 tests. Figure S1 summarizes the database construction and geocoding process.

#### 2.4 Metals measurements QA/QC

Tests that were analyzed at the State Inorganic Lab and were likely from schools or other community centers testing for lead were removed. Lastly, data were examined for extreme outliers of any individual metal and n=6 observations were removed. Specifically, two samples were removed because they indicated that they were “solid samples” and were therefore unlikely well water tests, two samples had implausibly high iAs (>1,200 ppb), one sample had implausibly high CrVI (>1,900 ppb), and one was removed for assessing iAs species rather than total iAs. Thus, n=117,960 tests were ultimately included in the NCWELL database.

The number of tests in which individual metals were evaluated varied and only metals with over 200 tests were included in subsequent statistical analysis and mapping (this removed tin, gold, and lithium measures). This resulted in a list of 28 metal/metalloids of interest included in the subsequent statistical analyses (Table S1). From these 28 metal/metalloids, different priority groupings of metals for subsequent analyses were defined according to relevance to public health and precision of data available. Specifically, 14 metals had at least one measurement in every county and thus were eligible for county-level analyses and from this set, six are toxic or have toxic properties (iAs, Cd, Pb, Mn, Hg, Cr), and three of these have previously been identified to be of particular concern to NC and are thus high priority contaminants (iAs, Mn and Pb) (Sanders et al. 2012, 2014; Tomlinson et al. 2019).

## 2.5 Statistical analysis of the NCWELL database

All statistical analysis was conducted in R (v4.0.2). For all metal concentrations reported as below the LOR, measurements were imputed as  $<di>LOR/ 2$ , as previously done (Rager et al. 2014, 2021). For all 28 metals of interest, descriptive statistics for the entire state were calculated. Descriptive statistics were calculated for all 100 counties and all census tracts where there was at least one test per tract (1,948 census tracts, or 89% of the 2,195 census tracts). Descriptive statistics included the number and percentage of reports above the EPA and NCGW standards. Many different types of standards are relevant to assessing the quality of well water in NC (Table S2). There are two of sets of federal level standards used for public water systems: (1) the enforceable (for public drinking water systems) National Primary Drinking Water Regulations (or EPA MCLs) (US EPA) and (2) the non-enforceable National Secondary Drinking Water Regulations (or EPA secondary MCL, (SMCL)), generally set based on odor and taste standards (US EPA). Rather than an MCL or SMCL, Pb is regulated by the EPA Action Level (AL) governed by the Lead and Copper Rule of 1991 (US EPA). The other sets of standards are NC-specific: (3) the NC Groundwater Standards (NCGW) (NC DEQ) (4) the NC Interim Maximum Allowable Concentration (NC IMAC) (NCDEQ) and lastly (5) the Coal Ash Management Act of 2014 (CAMA) established health protective screening levels (General Assembly of North Carolina 2013). In addition, Mn and CrVI have health advisory limits (HAL). Mn has an EPA HAL that is more directly relevant to public health than the SMCL and will thus be the standard of focus for subsequent analyses and mapping.

**2.5.1. Identifying counties of concern.**—To aid with public health prioritization of counties in NC, for three metals of particular concern to NC – iAs, Pb and Mn – counties were ranked according to the population-at-risk. The population-at-risk rank was determined by multiplying the percentage of wells exceeding the EPA MCL for iAs, EPA lifetime HAL for Mn or EPA AL for Pb, and the percentage of population within a county using private wells and ranking counties from the highest to the lowest. While the percentage of the population using private wells captures the proportion of the county that may be at risk from well water contamination without over weighting counties that are more populated, using the absolute number of the individuals using private wells may also be useful metric for prioritization. Further, the ranking of counties based simply on the percentage of tests that exceed the standard may be helpful in some cases as well. Therefore, we also calculated a number-at-risk rank following the same formula as the population-at-risk rank but utilizing the absolute number of the individuals on private wells, rather than the percentage of the population, and an exceedance-of-standard rank, based on the percentage of tests that exceed the standard and not including any information on the county population.

The number of individuals using private wells in a county was obtained from the US Geological Survey database for 2015 and the population of the county was obtained from the US Census Bureau Population Estimates for 2018 through the tidycensus R package (v 0.11.4) (US Geological Survey; Walker and Herman 2021). To visualize the counties of concern, for all six metals with toxic properties (iAs, Cd, Mn, Pb, Cr, Hg), maps at a county and census tract-level were generated utilizing the R packages ggplot2 (v 3.3.3),

tmap (v 3.3–1), tmaptools (v 3.3–1), biscale (v 0.20) and classInt (v 0.4–3) (Wickham 2016; Tennekes 2018, 2021; Bivand 2020; Prener et al. 2020).

### 2.5.2 Evaluating correlation and co-occurrence utilizing k-means analysis.—

To assess a basic measure of co-occurrence of metals, Spearman correlations between the county-level mean concentrations were calculated and visualized for all metals with at least one well water test report in all 100 counties (n=14). While the threshold was at least one well water test per county, the minimum number of tests that was averaged to calculate a county-level mean was 19 and the maximum was 7262. To assess if the co-occurrence also appeared at an individual well level, Spearman correlations were calculated for these metals at an individual well level as well. The Hmisc (v 4.5–0) and corrplot (0.88) R packages were utilized to calculate correlations (Harrell Jr. 2021; Wei and Simko 2021). Significant correlation was defined as  $p < 0.05$ .

Next, to construct clusters of counties with distinct profiles of metal concentrations, k-means was utilized, modelled after a framework previously used to geospatially cluster environmental exposures (Austin et al. 2013). The objective of this approach is to maximize the similarity of metal concentration profiles of iAs, Mn, Pb, Cd, Cr, and Hg within clusters, while maximizing differences between clusters. These metals were selected because they are either toxic or have toxic properties so that the interpretation of the clustering could focus on toxic metals of public health concern. The k-means algorithm seeks to partition M points (in this case, 100 counties) in N dimensions into k clusters (Hartigan and Wong 1979). In this application, N represents county-level mean concentrations across all years of data collection for the six metals listed above. The number of clusters (k) must be *a priori* selected based on either pre-existing knowledge or through metrics assessing observable responses of the data to the clustering. To maximize interpretability of the clustering solutions, we set the *a priori* preferences for the optimal k to be greater than 2 and less than 6 (ie. k=3, 4 or 5); however, we iterated through a larger range of k to observe trends in responses of the data to the clustering. To determine the optimal k for this dataset, we iterated through k=1 to k=99 and evaluated four metrics that captured the balance between the difference between clusters and the similarity within clusters, while also considering the interpretability preferences. The four metrics were evaluated sequentially and hierarchically, with each metric dictating possible solutions evaluated with the next metric: (1) the elbow point; (2) the number of single county clusters; (3) the gain in significant differences in metal concentrations between counties; (4) the majority rule of 30 different indices designed to determine the optimal k, generated through the Nbclust function within the factoextra package (v1.07) (Kassambara and Mundt 2020). Further details regarding evaluations of the aforementioned metrics is provided in Supplementary Information, Section 2.5. As two sensitivity analyses, we reran the entire k-means analysis, including calculating the metrics described above, for two additional scenarios. First, removing Cd from the included metals because of its skewed state-wide distributions which could result in it disproportionately influencing the clustering results. Second, removing the maximum individual measurement of Pb as this could potentially skew the county (Tyrell) mean significantly and thus disproportionately influence the clustering results.



**2.5.3 Examining temporal trends.**—To statistically test trends in time, for each county, for iAs, Cd, Pb, Mn, Hg and Cr, simple linear regression models were run evaluating yearly differences in the county-level percentage of exceedances of the EPA standard (MCL for iAs, Cd, Hg and Cr, HAL for Mn, AL for Pb). The independent variable was the year (range: 1998–2019) and the dependent variable was the yearly county-level percentage of exceedances. A significant temporal trend was defined as Benjamini-Hochberg adjusted p-value <0.05 (Benjamini and Hochberg 1995). To visually examine changes in metal concentrations over time for the metals of particular interest – iAs, Mn and Pb – heatmaps were generated representing the percentage of exceedances of the EPA MCL, AL or HAL for each year of data collection for each of the top 25 counties, as defined in the counties of concern analysis described above.

## 2.6 NCWELL database availability statement

The NCWELL database is publicly available on the UNC Superfund Research Program UNC Dataverse, <https://doi.org/10.15139/S3/BDQG90> (Eaves et al. 2021). Specifically, three relevant datasets with associated metadata are accessible. These datasets are Dataset #1: County-level distributions of the included 28 metals/metalloids as well as nitrites and nitrates; Dataset #2: Census tract-level distributions of the included 28 metals/metalloids as well as nitrites and nitrates; Dataset #3: Relevant standards and screening levels for metals/metalloids in NC groundwater as of July 2021. To protect the privacy of private well owners, the dataset that includes individual well level data, including their point locations, is not publicly available online. These data can be made available upon direct request to the authors.

## 3. Results

### 3.1 Generation of the NCWELL database.

The database contained a total of n=117,960 geocoded well water tests representing the timeframe between October 19, 1998, and May 20, 2019, with test reports from all 100 counties and 89% of census tracts. Of n=117,960 total well water tests, n=93,913 (79.61%) had unique latitude and longitude coordinates, likely representing unique wells. Based on identical latitude and longitude coordinates, n=12,989 wells were estimated to have more than one test included in the dataset. While challenging to estimate the proportion of all wells in the state that are included in this database as the exact number of wells in NC is not known, we did assess the relationship between the estimated population relying on private well water (based on US Geological Survey estimates for 2015) and the number of tests for iAs, Mn and Pb (US Geological Survey). There was a strong linear correlation between the number of tests and the population relying on well water per county (Spearman rho = 0.75, p-value <0.001, for each of the metals).

The pre-2010 data that was previously published included only iAs, Pb, Cd and Mn, therefore these metals have a larger number of included well water tests in the database (approximately n=110,000 for iAs, Pb, Mn and n=75,393 for Cd) (Sanders et al. 2012, 2014) (Table S3). In addition, many other metals have over 65,000 well water tests in which they were measured, namely, in descending order of number of well water tests, Cu, Fe, Mg, Cr,

Ca, Ba, Se, Zn, Ag, Na (Table S2). The remaining metals, including metals of concern – U, CrVI, V — were each measured in less than 1% of the collected samples (Table S2).

### 3.2 Private well water metal concentrations exceed federal and state standards.

Of 110,383 included tests for iAs, 2,502 (2.3%) exceeded the EPA MCL, 10ppb. Of 110,182 included tests for Mn, 26,816 (24.34%) exceeded the EPA SMCL, 50ppb, and 5,409 (4.9%) exceeded the EPA HAL, 300 ppb. Lastly, of 110,409 included tests for Pb, 3076 (2.8%) exceeded the EPA AL, 15 ppb (Table 1). Alarming high maximum reported levels were found for high priority contaminants: iAs (806 ppb), Mn (46,300 ppb) and Pb (105,440 ppb).

### 3.3 The population-at-risk rank identifies counties of concern for high priority contaminants, iAs, Mn and Pb.

When evaluating county-level variation in metal concentrations in private wells, geospatial differences were observed across all metals of interest (Figure 1, Figure S2). To identify counties of concern, the population-at-risk rank, which incorporated both the percentage of exceedances of standards and the percentage of the population relying on private wells, was calculated for each county, as in our prior study (Sanders et al. 2012). For iAs, the top ten counties according to population-at-risk are: Anson, Stanly, Union, Alexander, Randolph, Lincoln, Person, Transylvania, Rockingham, and Chatham. In Anson, Stanly, and Union counties, over 18% of well water tests exceed the EPA and NCGW standard (Table 2, Figure 1). For Mn, the top ten counties according to population-at-risk are: Person, Caswell, Stanly, Granville, Chatham, Anson, Clay, Perquimans, Wilson, Randolph (Table 2, Figure 1). The county with the highest percentage of well water tests exceeding the HAL for Mn was Perquimans County (21.2%). For Pb, the top ten counties according to population-at-risk are: Alleghany, Madison, Macon, Ashe, Surry, McDowell, Caswell, Hoke, Yancey, and Person (Table 2, Figure 1). The county with the highest percentage of well water tests exceeding the AL for Pb was Hoke County (12.8%).

In addition to the population-at-risk rank, a number-at-risk rank, which utilizes the absolute number of individuals using private well within a county rather than the percentage, and an exceedance-of-standard rank for each county can be found in Table S5. County-level variations in exceedances and mean-levels for Cd, Cr and Hg are described in Table S5 and Figure S2 and S3. Census tract-level maps and descriptive statistics were also generated for iAs, Mn, Pb, Cd, Cr and Hg (Table S6, Figure S4&5).

### 3.4 Correlation of key metal pairings is common.

Numerous metal pairings were correlated in pairwise Spearman correlation assessments of county-level mean concentrations (Table S7, Figure 2). Most correlations were positive, suggestive of co-occurrence. The two most significantly correlated pairs of metals that include a toxic metal were iAs and Mn (coefficient = 0.29, p-value=0.003) and Cu and Pb (coefficient = 0.53, p-value < 0.001) and (Table S7, Figure 3). These two pairings were also significantly positively correlated at an individual well level. Across all metal pairs, there was a general convergence of the direction and significance of correlations in both the county-level and individual well analyses (Table S7).

### 3.5 Four distinct clusters of counties were identified using k-means clustering.

K means clustering was used to identify four clusters ( $k=4$ ). Figure S6 shows (A) the change in within cluster variation with the addition of clusters used to determine the elbow point and (B) the number of single county clusters with the addition on clusters. While the elbow point was located at approximately  $k=7/8$ ,  $k=4$  still represented a significant decrease (~45%) in the proportion of within cluster variation of the total variation. There are also no single county clusters with  $k=4$ . Additional metrics are detailed in Table S8. There were no further gains in the number of metals with significant differences in means between clusters after  $k=4$  (comparing  $k=3, 4$  and  $5$ ). For the multi-indices' summary measure,  $k=6$  had the highest number of indices, followed by  $k=4$ . Additionally, sensitivity assessments did not change the metrics for the optimal  $k$  substantially (Table S8).

The four clusters represent four general patterns of toxic metal co-occurrence in counties in NC (Table S5, Figure 4). To detail, Cluster #1 contains the largest number of counties ( $n=66$ ) and represents counties with generally low levels of metal concentrations in well water. For each metal, the mean of the z-score standardized mean concentration for each county in Cluster #1 was below the state-wide mean (Figure 4B). Cluster #2, comprising 28 counties, is influenced primarily by higher levels of Hg and low levels of all other toxic metals, although this does appear to be driven primarily by a couple counties (Figure 4B, 4C). Cluster #1 and Cluster #2 are termed “low-metal” clusters.

Cluster #3 contains four counties (Guildford, Mecklenburg, Tyrell, Wake) and is distinguished primarily through the high concentrations of Pb observed and will thus be termed “Pb-priority” cluster (Table S5, Figure 4B, 4C). While Cd also appears high in this cluster, this should be interpreted cautiously given the skewed distribution of Cd across the state, and because counties in this cluster have only a very small (<0.5%) shift in the mean of the percentage of exceedances of the EPA MCL compared to the other clusters (Figure 4C). Cluster #4 contains 3 counties (Anson, Union, Stanly) with distinctly high levels of iAs and Mn, thus will be termed the “iAs-Mn-priority” cluster. In fact, the percentage of exceedances of the EPA MCL for iAs for counties in this cluster are all over 18%, while all other counties have less than 10% of tests exceeding the standard (Table 2, Figure 4C). Similarly, for Mn, the average of the percentage of tests exceeding the EPA HAL is markedly higher than for all other counties (Figure 4C). The iAs-Mn-priority cluster is tightly distributed around the mid-South region of the state while Pb-priority cluster is not tightly geospatially clustered and generally is comprised of urban and peri-urban counties (Figure 4A). Comparing the mean levels of metals across the clusters, for all metals other than Cr, the clusters differed significantly (Table 3).

### 3.6 There is minimal temporal variation in exceedances of standards at a county-level.

State-wide yearly distributions for year 1998–2019 are detailed for all 28 metals of interest in Table S9. Because only iAs, Mn, Pb and Cd data were available pre-2010, the other metals have zero well water tests until 2010. There is a general trend of an increasing number of tests after 2008, which is expected given the change in the law in 2008 to require new wells be tested (NC Department of Public Health, Epidemiology section). Before 2008, for iAs, Mn and Pb approximately 3000–4500 tests were conducted each year, then after

2008, 6000–7500 tests were conducted per year, with minimal variation within the pre-/post-2008 time periods. To examine whether the percentage of tests exceeding standards were relatively stable over time, simple linear regression models assessing the relationship between year (independent variable) and percentage of exceedances of the EPA MCL/SMCL/HAL/AL (dependent variable) were run for iAs, Cd, Pb, Mn, Hg and Cr for each county. Highlighting their sustained nature over time, no metals demonstrated a significant (BH-adjusted  $p < 0.05$ ) linear time trend in relation to the percentage of exceedances within each county (Table S10). Additionally, for iAs, Mn and Pb, this pattern of minimal temporal variation appears to hold for the top 25 counties of concern (Figure 5).

## 4 Discussion

A growing body of research documents the environmental health challenge of metal exposure via private well water in NC. For example, high levels of iAs and Mn have been associated with increased risk of children being born with birth defects (Sanders et al. 2012, 2014). Additionally, documented disparities in Pb blood levels exist between well users and non-well users (Gibson et al. 2020). Moreover, there are concerns about CrVI and V contamination across the state (Tomlinson et al. 2019; Coyte et al. 2020). Despite these concerns, to date there has been no updated, systematic evaluation of private well water metal concentrations in NC. The present study fills this gap by establishing the NCWELL database, a dataset of approximately 118,000 well water tests over a twenty-year period. The database includes 28 different metal/metalloids, summarized at the census tract and county-level. In the present study, we also present two public health prioritization tools for addressing metal levels in well water in NC. First, we generated a population-at-risk rank for iAs, Mn and Pb, a metric that combines the proportion of a county on well water and the percentage of exceedances of the EPA standard or health advisory level, for all counties. Second, we identified a cluster of counties with distinctly high iAs and Mn (iAs-Mn-priority cluster) and a cluster with high levels of Pb (Pb-priority cluster). It is anticipated that these data will serve as a resource for public health professionals, researchers, and NC residents to identify communities at risk where interventions to minimize exposure are most critical. These results from the present study can be used by local health departments and other invested stakeholders to identify priority counties of concern as either those with high potential for joint exposure to multiple metals or a high population-at-risk rank for any single metal.

A striking finding of this study was the extent of high levels of iAs and Mn contamination in private wells and their strong pattern of co-occurrence. More than 2,500 tests exceeded 10 ppb (EPA MCL) for iAs and more than 5,000 tests exceeded 300 ppb (EPA HAL) for Mn. Of note, three counties – Anson, Union, and Stanly – were identified as comprising the iAs-Mn priority cluster where joint exposure to high levels of iAs and Mn is likely. These counties were also identified as the top three counties by the population-at-risk rank for iAs. In each of these counties, the percentage of tests exceeding the iAs standard was over 18%, while the percentage exceeding the standard across the entire state was 2.3%. In our previous assessment of iAs in NC wells, Anson, Union, and Stanly were also the highest ranked counties for exceeding standards and population-at-risk ranking (Sanders et al. 2012). The top three counties of concern for Mn identified through the population-at-risk

rank included Person, Caswell, and Stanly counties. Union and Anson were also included in overall top 25 population-at-risk counties for Mn, underscoring the southern Piedmont region of NC as an area of concern for joint exposure to iAs and Mn. As we observed no linear trend in the percentage of exceedances of the EPA MCL/HAL for iAs or Mn over time for any county, it is likely that these regions of the state have had consistently high levels of iAs and Mn over the documented twenty-year period. This supports previous research demonstrating that the Carolina terrane, the bedrock beneath most of the Piedmont of NC, is likely the major geogenic contributor to iAs and Mn in private wells (Pippin et al. 2003; Kim et al. 2011; Sanders et al. 2012; Gillispie et al. 2016).

In addition to iAs and Mn, high levels of Pb were observed in the NCWELL database. Specifically, 2.8% of wells tested (> 3000) had levels of Pb above the EPA AL. Perhaps more importantly given the effects of even low-dose Pb, approximately 8% (approximately 9000) of wells tested had detectable levels of Pb. The Pb-priority cluster included Guilford, Mecklenburg, Tyrell, and Wake counties, likely distinct from other counties due to wells being in peri-urban areas, with older homes where the potential for premise plumbing corrosion is high. The population-at-risk rank identified Alleghany, Madison, and Macon counties as the top 3 counties of concern for Pb in well water. Importantly, the Pb findings in the current study will be underestimates of well water Pb levels as the samples collected were flush samples, which have been shown to reduce the concentration of Pb significantly compared to a first draw (Pieper et al. 2015, 2018). Further, many tests are conducted at the well head and thus do not represent lead contamination from possible premise plumbing.

NC residents relying on private well water are provided limited support and face many barriers to mitigating their potential exposure to metals via well water. The median cost of a well water test for inorganic analytes in NC is \$90 and this cost barrier has been identified as a significant impediment to regular testing (Stillo et al. 2019; Gibson et al. 2020; Wait et al. 2020). Furthermore, costs to treat a contaminated well are even higher: even basic kitchen tap filter may cost up to \$200, under-sink filters may be \$300–400, and require replacement cartridges of \$50 twice a year (Gibson et al. 2020). Additionally, evidence regarding how effective these filters are after long-term use is limited, although an approximately \$100 filter did remain effective at removing Pb after six months of use in a recent NC-based study (Mulhern and MacDonald Gibson 2020). Other solutions are often even more expensive and can include well depth modification or use of an alternative water source such as bottled water (Pratson et al. 2010; Sanders et al. 2012). Of course, these costs are prohibitive for many low-income families, resulting in socioeconomic disparities in testing and treatment and in turn, disparities in actual exposure from contaminated well water, highlighting the environmental justice concerns relating to well water in NC (Flanagan et al. 2016).

In fact, viewing these findings with an environmental justice lens yields other important insights. Municipal underbounding is the practice of borders growing to engulf poor and minority communities without expanding municipal services such as water and sewer lines (Aiken 1987; Leker and MacDonald Gibson 2018). A study based in Wake county, NC found that for every 10% increase in the African American population proportion within a census block, the odds of exclusion from municipal water service increased by 3.8%, clear evidence of environmental racism (MacDonald Gibson et al. 2014; Nigra 2020). Without

public water supply, residents rely on private well water and as such, underbounding enhances racial disparities in well-water related exposures. One of the recommendations from the 2015 N.C. Research Triangle Environmental Health Collaborative “Safe Water from Every Tap” summit was to provide additional support and resources to individuals for monitoring and maintenance of their private wells (Gibson and Pieper 2017). We hope that the NCWELL database will be a resource for residents and stakeholders for this very purpose. Furthermore, the striking findings regarding high levels of metals in NC contribute to a growing literature highlighting the need for immediate public health interventions to reduce exposures for households reliant on private wells.

While this study represents the generation of one of the largest repositories of private well water tests in NC to date, it is not without limitations. The NCWELL database does not represent all wells in NC, only wells that owners had tested in the data collection period. As noted above, the likelihood of testing is greatly influenced by socioeconomic factors therefore it likely underestimates the environmental justice concerns that are present. The database also does not contain tests done through private companies, and thus the current data may underestimate actual exposure scenarios. Further, the database does contain some repeated tests on the same well, which could inflate the metrics, such as percentage of exceedances, calculated. Further, while these tests are reported below the LOR, metals/metalloids may still be present at low concentrations that can still be hazardous to human health and this low-level contamination is not captured in the NCWELL database. Lastly, this analysis focused on county-level assessment; however, variability in concentrations does exist within counties as demonstrated by the census-tract level maps provided. We encourage other users to conduct census-tract level or smaller analyses to assess this variability.

## 5. Conclusions

In conclusion, the NCWELL database represents a novel resource for researchers, environmental health professionals, and residents to conduct monitoring and examine trends in well water metal concentrations. Alarming high levels of iAs, Pb and Mn were observed in the NCWELL database which should confirm that exposure to toxic metals via well water remains a critical public health concern in NC. This study has highlighted the need to take a mixtures-based approach to tackle well water-based metal exposure by identifying specific clusters of counties with multi-metal contamination. Overall, this study underscores the need for universal monitoring and screening of private wells as well as focused interventions to minimize toxic exposure and protect public health of NC residents.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank Dr. Crystal Lee Pow Jackson for her insights and comments on the draft of this manuscript. We are also grateful for Phillip McDaniel’s guidance on geocoding the well water data. In addition, any reader interested in further guidance and support regarding private wells, including recommended testing and contact information for the private well water program at the DHHS, can find information at: 1) <https://>

[epi.dph.ncdhhs.gov/oe/programs/wellwater.html](https://epi.dph.ncdhhs.gov/oe/programs/wellwater.html); 2) <https://epi.dph.ncdhhs.gov/oe/wellwater/whentotest.html>; 3) <https://ehs.ncpublichealth.com/oswp/wells-resources.htm>.

### Funding sources

This work was supported by the National Institute of Environmental Health Sciences (1R01ES029531, P42ES031007).

### Bibliography

- Aiken CS. Race as a factor in municipal underbounding. *Annals of the Association of American Geographers*. 1987 Dec;77(4):564–79.
- Andrews KW, Savitz DA, Hertz-Picciotto I. Prenatal lead exposure in relation to gestational age and birth weight: a review of epidemiologic studies. *Am J Ind Med*. 1994 Jul;26(1):13–32. [PubMed: 8074121]
- ATSDR. Substance Priority List. ATSDR; 2019.
- Austin E, Coull BA, Zanobetti A, Koutrakis P. A framework to spatially cluster air pollution monitoring sites in US based on the PM<sub>2.5</sub> composition. *Environ Int*. 2013 Sep;59:244–54. [PubMed: 23850585]
- Ayotte JD, Montgomery DL, Flanagan SM, Robinson KW. Arsenic in groundwater in eastern New England: occurrence, controls, and human health implications. *Environ Sci Technol*. 2003 May 15;37(10):2075–83. [PubMed: 12785510]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995 Jan;57(1):289–300.
- Bivand R. classInt: Choose Univariate Class Intervals. <https://CRAN.R-project.org/package=classInt>; 2020.
- Coyte RM, McKinley KL, Jiang S, Karr J, Dwyer GS, Keyworth AJ, et al. Occurrence and distribution of hexavalent chromium in groundwater from North Carolina, USA. *Sci Total Environ*. 2020 Apr 1;711:135135.
- Dieter CA, Maupin MA, Caldwell RR, Harris MA, Ivahnenko TI, Lovelace JK, et al. Estimated use of water in the United States in 2015. USGS Circular 1441; 2018.
- Eaves LA, Keil A, Henry A, Fry RC. A dataset describing well water metal/metalloid contamination of private wells in North Carolina, 1998–2019 [Internet]. 2021. Available from: 10.15139/S3/BDQG90
- Ferguson KK, Chin HB. Environmental chemicals and preterm birth: Biological mechanisms and the state of the science. *Curr Epidemiol Rep*. 2017 Mar;4(1):56–71. [PubMed: 28944158]
- Flanagan SV, Spayd SE, Procopio NA, Marvinney RG, Smith AE, Chillrud SN, et al. Arsenic in private well water part 3 of 3: Socioeconomic vulnerability to exposure in Maine and New Jersey. *Sci Total Environ*. 2016 Aug 15;562:1019–30. [PubMed: 27118035]
- General Assembly of North Carolina. Senate Bill 729: Coal Ash Management Act. 729 2013.
- Gibson JM, Fisher M, Clonch A, MacDonald JM, Cook PJ. Children drinking private well water have higher blood lead than those with city water. *Proc Natl Acad Sci USA*. 2020 Jul 21;117(29):16898–907. [PubMed: 32631989]
- Gibson JM, Pieper KJ. Strategies to Improve Private-Well Water Quality: A North Carolina Perspective. *Environ Health Perspect*. 2017 Jul 7;125(7):076001.
- Gillispie EC, Austin RE, Rivera NA, Bolich R, Duckworth OW, Bradley P, et al. Soil weathering as an engine for manganese contamination of well water. *Environ Sci Technol*. 2016 Sep 20;50(18):9963–71. [PubMed: 27570123]
- Harrell FE Jr. Hmisc: Harrell Miscellaneous. <https://CRAN.R-project.org/package=Hmisc>; 2021.
- Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. *Appl Stat*. 1979;28(1):100–8.
- Kassambara A, Mundt F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. <https://CRAN.R-project.org/package=factoextra>; 2020.

- Kim D, Miranda ML, Tootoo J, Bradley P, Gelfand AE. Spatial modeling for groundwater arsenic levels in North Carolina. *Environ Sci Technol*. 2011 Jun 1;45(11):4824–31. [PubMed: 21528844]
- Kim M-J, Nriagu J, Haack S. Arsenic species and chemistry in groundwater of southeast Michigan. *Environ Pollut*. 2002;120(2):379–90. [PubMed: 12395852]
- Leker HG, MacDonald Gibson J. Relationship between race and community water and sewer service in North Carolina, USA. *PLoS One*. 2018 Mar 21;13(3):e0193225.
- Levallois P, Barn P, Valcke M, Gauvin D, Kosatsky T. Public health consequences of lead in drinking water. *Curr Environ Health Rep*. 2018;5(2):255–62. [PubMed: 29556976]
- MacDonald Gibson J, DeFelice N, Sebastian D, Leker H. Racial disparities in access to community water supply service in Wake County, North Carolina. *Am J Public Health*. 2014 Dec;104(12):e45–e45. [PubMed: 24228652]
- Mecklenburg County Health Department. Mecklenburg County Health Dept. Well Information System 4.0 [Internet]. [cited 2019 Apr 30]. Available from: <https://edmsmapserver.mecklenburgcountync.gov/wis4/>
- Mulhern R, MacDonald Gibson J. Under-Sink Activated Carbon Water Filters Effectively Remove Lead from Private Well Water for over Six Months. *Water (Basel)*. 2020 Dec 21;12(12):3584.
- Naujokas MF, Anderson B, Ahsan H, Aposhian HV, Graziano JH, Thompson C, et al. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environ Health Perspect*. 2013 Mar;121(3):295–302. [PubMed: 23458756]
- NC Department of Public Health, Epidemiology section. Well Water Fact Sheet [Internet]. [cited 2021a Oct 5]. Available from: <https://epi.dph.ncdhhs.gov/oe/doc/PrivateWellWaterFactST.pdf>
- NC Department of Public Health, Epidemiology section. Well Water FAQ [Internet]. [cited 2021b May 17]. Available from: <https://epi.dph.ncdhhs.gov/oe/wellwater/whentotest.html>
- NC DEQ. NC Groundwater Standards (Substance 15A NCAC 02L.0202) [Internet]. [cited 2020 Aug 21]. Available from: <https://deq.nc.gov/documents/nc-stds-groundwater02lstandards>
- NC Division of Public Health. Search for Inorganic Chemistry report [Internet]. [cited 2021 Oct 5]. Available from: <https://celr.ncpublichealth.com/InOrganicChemistry>
- NCDEQ. NC Interim Maximum Allowable Concentrations (IMACs) [Internet]. [cited 2021 Jul 5]. Available from: [https://files.nc.gov/ncdeq/Water%20Quality/Planning/CSU/Ground%20Water/APPENDIX\\_I\\_IMAC\\_2-01-21.pdf](https://files.nc.gov/ncdeq/Water%20Quality/Planning/CSU/Ground%20Water/APPENDIX_I_IMAC_2-01-21.pdf)
- Needleman HL, Bellinger D. The health effects of low level exposure to lead. *Annu Rev Public Health*. 1991;12:111–40. [PubMed: 1828669]
- Nigra AE. Environmental racism and the need for private well protections. *Proc Natl Acad Sci USA*. 2020 Jul 28;117(30):17476–8. [PubMed: 32641505]
- Nigra AE, Sanchez TR, Nachman KE, Harvey D, Chillrud SN, Graziano JH, et al. 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 Planet Health*. 2017;2(11):e513–21.
- O’Neal SL, Zheng W. Manganese toxicity upon overexposure: a decade in review. *Curr Environ Health Rep*. 2015 Sep;2(3):315–28. [PubMed: 26231508]
- Ooms J. pdftools: Text Extraction, Rendering and Converting of PDF Documents. <https://CRAN.R-project.org/package=pdfutils>; 2021.
- Peters SC, Blum JD, Klaue B, Karagas MR. Arsenic occurrence in new hampshire drinking water. *Environ Sci Technol*. 1999 May;33(9):1328–33.
- Pieper KJ, Krometis L-AH, Gallagher DL, Benham BL, Edwards M. Incidence of waterborne lead in private drinking water systems in Virginia. *J Water Health*. 2015 Sep;13(3):897–908. [PubMed: 26322775]
- Pieper KJ, Nystrom VE, Parks J, Jennings K, Faircloth H, Morgan JB, et al. Elevated lead in water of private wells poses health risks: case study in macon county, north carolina. *Environ Sci Technol*. 2018 Apr 3;52(7):4350–7. [PubMed: 29536726]
- Pippin C, Butczynski M, Clayton J. Distribution of Total Arsenic in Groundwater in the North Carolina Piedmont. NC Department of Environment and Natural Resources; 2003.



- Poropat AE, Laidlaw MAS, Lanphear B, Ball A, Mielke HW. Blood lead and preeclampsia: A meta-analysis and review of implications. *Environ Res.* 2018;160:12–9. [PubMed: 28938191]
- Pratson E, Vengosh A, Dwyer G, Pratson L, Klein E. The Effectiveness of Arsenic Remediation from Groundwater in a Private Home. *Ground Water Monit Remediat.* 2010 Feb;30(1):87–93.
- Prener C, Grossenbacher T, Zehr A. biscale: Tools and Palettes for Bivariate Thematic Mapping. <https://CRAN.R-project.org/package=biscale>; 2020.
- Rager JE, Bailey KA, Smeester L, Miller SK, Parker JS, Laine JE, et al. Prenatal arsenic exposure and the epigenome: altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood. *Environ Mol Mutagen.* 2014 Apr;55(3):196–208. [PubMed: 24327377]
- Rager JE, Clark J, Eaves LA, Avula V, Niehoff NM, Kim YH, et al. Mixtures modeling identifies chemical inducers versus repressors of toxicity associated with wildfire smoke. *Sci Total Environ.* 2021 Feb 10;775:145759.
- Sanders AP, Desrosiers TA, Warren JL, Herring AH, Enright D, Olshan AF, et al. Association between arsenic, cadmium, manganese, and lead levels in private wells and birth defects prevalence in North Carolina: a semi-ecologic study. *BMC Public Health.* 2014 Sep 15;14:955. [PubMed: 25224535]
- Sanders AP, Messier KP, Shehee M, Rudo K, Serre ML, Fry RC. Arsenic in North Carolina: public health implications. *Environ Int.* 2012 Jan;38(1):10–6. [PubMed: 21982028]
- Shaw WD, Walker M, Benson M. Treating and drinking well water in the presence of health risks from arsenic contamination: results from a U.S. hot spot. *Risk Anal.* 2005 Dec;25(6):1531–43. [PubMed: 16506980]
- Spangler AH, Spangler JG. Groundwater manganese and infant mortality rate by county in North Carolina: an ecological analysis. *Ecohealth.* 2009 Dec;6(4):596–600. [PubMed: 20232227]
- Stillo F, Bruine de Bruin W, Zimmer C, Gibson JM. Well water testing in African-American communities without municipal infrastructure: Beliefs driving decisions. *Sci Total Environ.* 2019 Oct 10;686:1220–8. [PubMed: 31412518]
- Tennekes M. tmap: Thematic Maps in R. *Journal of Statistical Software*; 2018.
- Tennekes M. tmaptools: Thematic Map Tools. <https://CRAN.R-project.org/package=tmaptools>; 2021.
- Tomlinson MS, Bommarito P, George A, Yelton S, Cable P, Coyte R, et al. Assessment of inorganic contamination of private wells and demonstration of effective filter-based reduction: A pilot-study in Stokes County, North Carolina. *Environ Res.* 2019 Aug 2;177:108618.
- US EPA. Lead and Copper Rule [Internet]. [cited 2021a Oct 4]. Available from: <https://www.epa.gov/dwreginfo/lead-and-copper-rule>
- US EPA. National Primary Drinking Water Regulations | Ground Water and Drinking Water | US EPA [Internet]. [cited 2020b Apr 8]. Available from: <https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations>
- US EPA. Secondary Drinking Water Standards: Guidance for Nuisance Chemicals | Safe Drinking Water Act (SDWA) | US EPA [Internet]. [cited 2021c Jul 6]. Available from: <https://www.epa.gov/sdwa/secondary-drinking-water-standards-guidance-nuisance-chemicals>
- US Geological Survey. USGS Water Use Data for North Carolina for 2015 [Internet]. [cited 2021 Feb 1]. Available from: <https://waterdata.usgs.gov/nc/nwis/wu>
- Wait K, Katner A, Gallagher D, Edwards M, Mize W, Jackson CLP, et al. Disparities in well water outreach and assistance offered by local health departments: A North Carolina case study. *Sci Total Environ.* 2020 Dec 10;747:141173.
- Walker K, Herman M. tidy census: Load US Census Boundary and Attribute Data as “tidyverse” and ‘sf’-Ready Data Frames. <https://CRAN.R-project.org/package=tidy census>; 2021.
- Wei T, Simko V. “corrplot”: Visualization of a Correlation Matrix. <https://github.com/taiyun/corrplot>; 2021.
- Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag; 2016.
- Yang Q, Jung HB, Culbertson CW, Marvinney RG, Loiselle MC, Locke DB, et al. Spatial pattern of groundwater arsenic occurrence and association with bedrock geology in greater Augusta, Maine. *Environ Sci Technol.* 2009 Apr 15;43(8):2714–9. [PubMed: 19475939]

Zheng Y, Flanagan SV. The Case for Universal Screening of Private Well Water Quality in the U.S. and Testing Requirements to Achieve It: Evidence from Arsenic. *Environ Health Perspect.* 2017 Aug 3;125(8):085002.

Author Manuscript

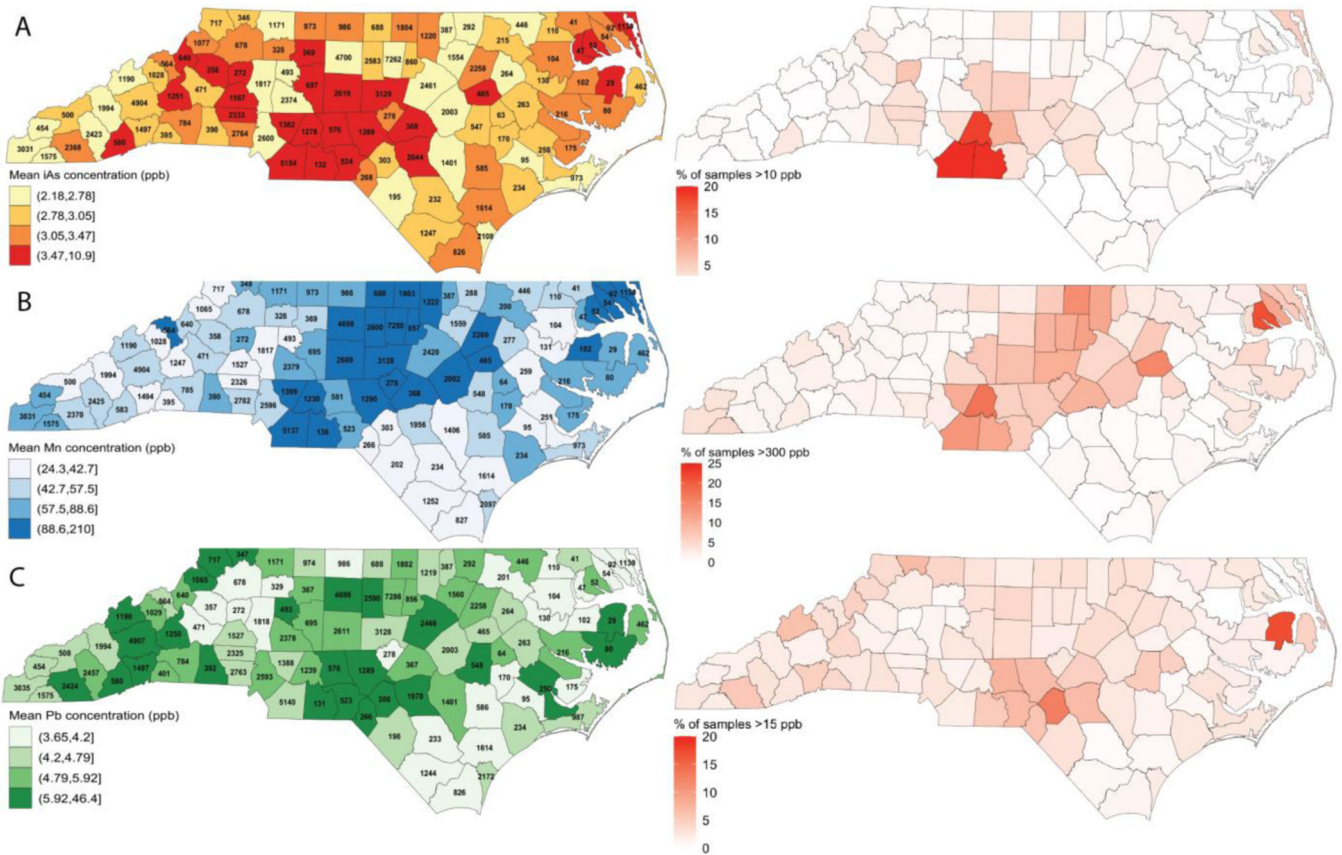
Author Manuscript

Author Manuscript

Author Manuscript

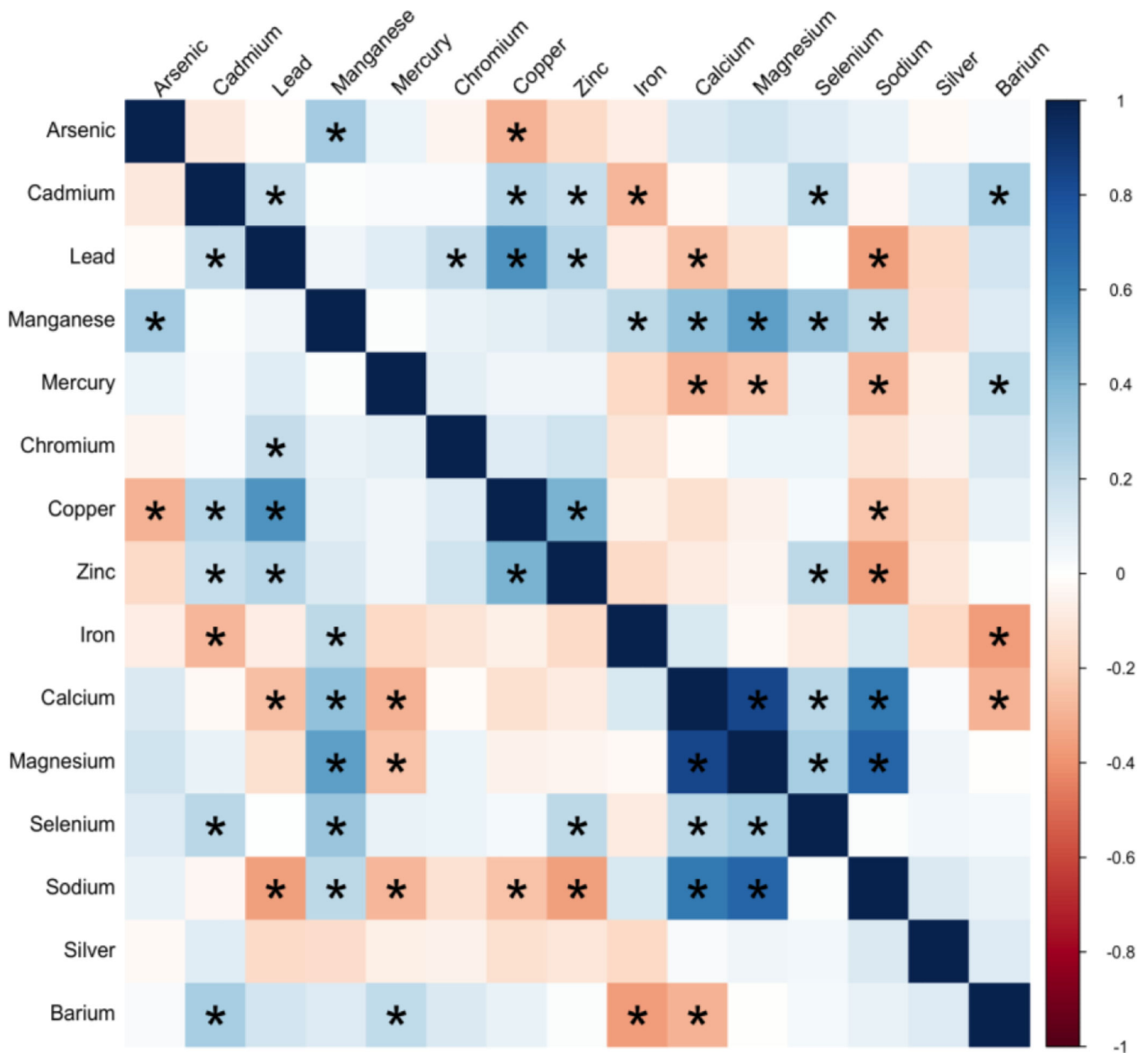
### Highlights

- NCWELL database constructed: n=117,960 well water tests over a 20-year period
- Thousands of well water tests exceeded EPA standards for iAs, Pb and Mn
- Two critical clusters of counties identified: (1) iAs-Mn cluster (2) Pb cluster
- Minimal temporal change in county-level percentage of exceedances of standards

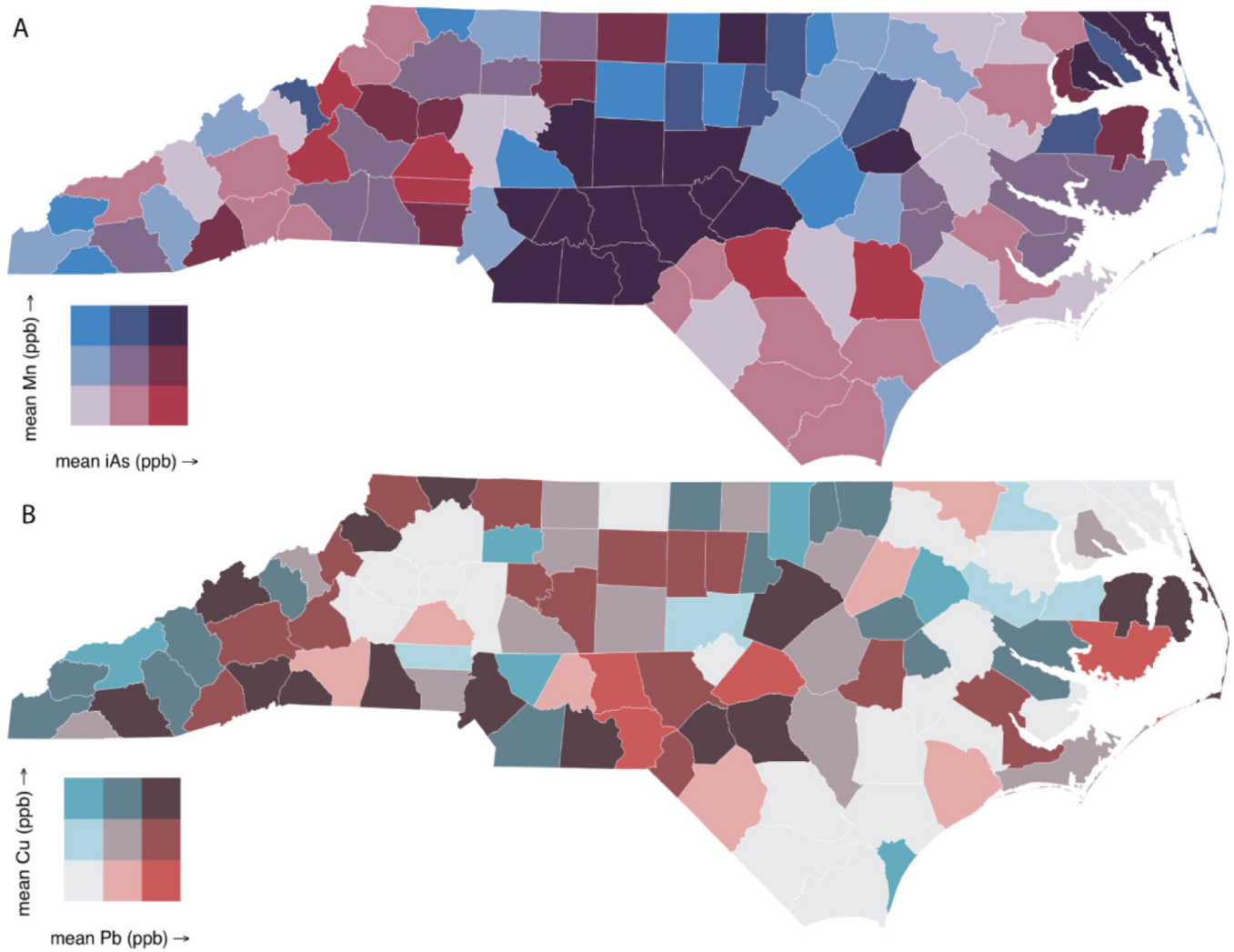


**Figure 1. Maps of North Carolina representing the average levels of (A) arsenic, (B) manganese and (C) lead within counties across North Carolina.**

The maps on the left represent the mean concentration within a county, based on quartiles (indicated by the color-shading), and the numbers in the county indicate the number of tests from that county, summarizing well water tests from October 19, 1998, to May 20, 2019. The maps on the right show the percentage of tests that recorded concentrations at or above the EPA MCL standard (for iAs and Pb) or the EPA Health Advisory Limit (for Mn) within a county, indicated by the color-shading. County-level and census-tract level maps of additional metals of particular interest (cadmium, chromium, copper, mercury, and zinc) can be found in Supporting Information.

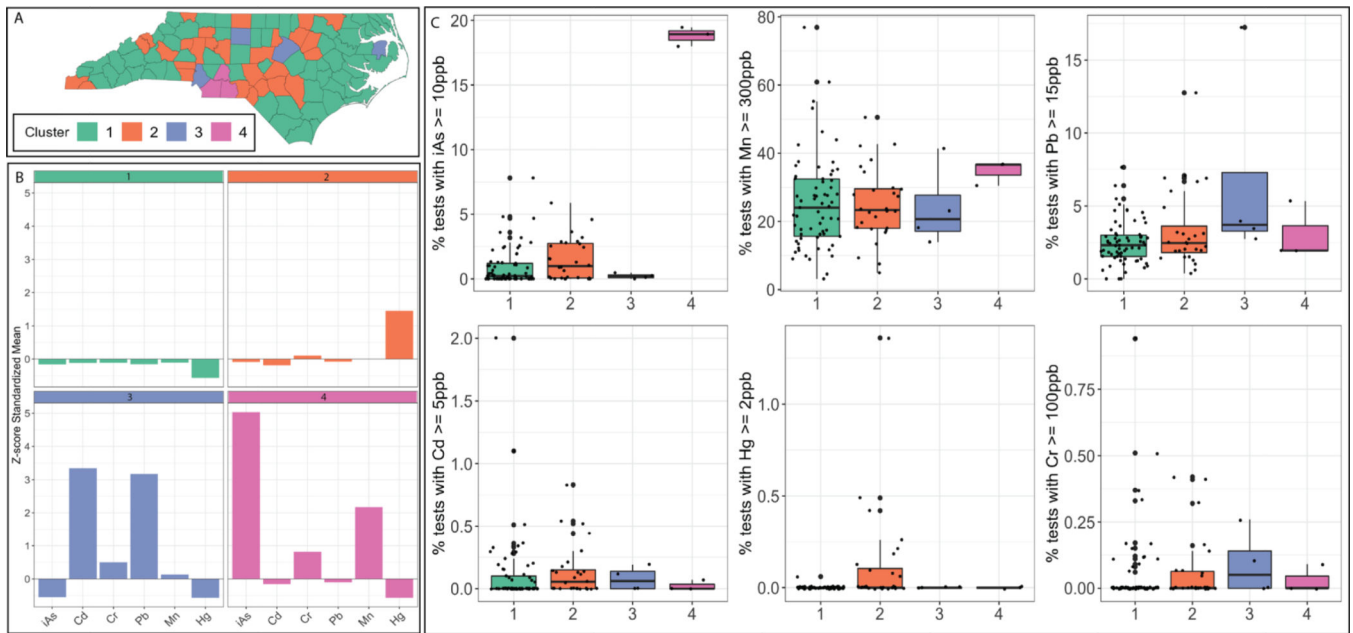


**Figure 2. Correlation between county mean metal levels.**  
 The color gradient indicates the Spearman correlation coefficient with blue indicating a positive association. Significant ( $p < 0.05$ ) correlations are indicated with an \* ( $p < 0.05$ ). All metals with at least one well water test for all 100 counties were included in the correlation analysis.

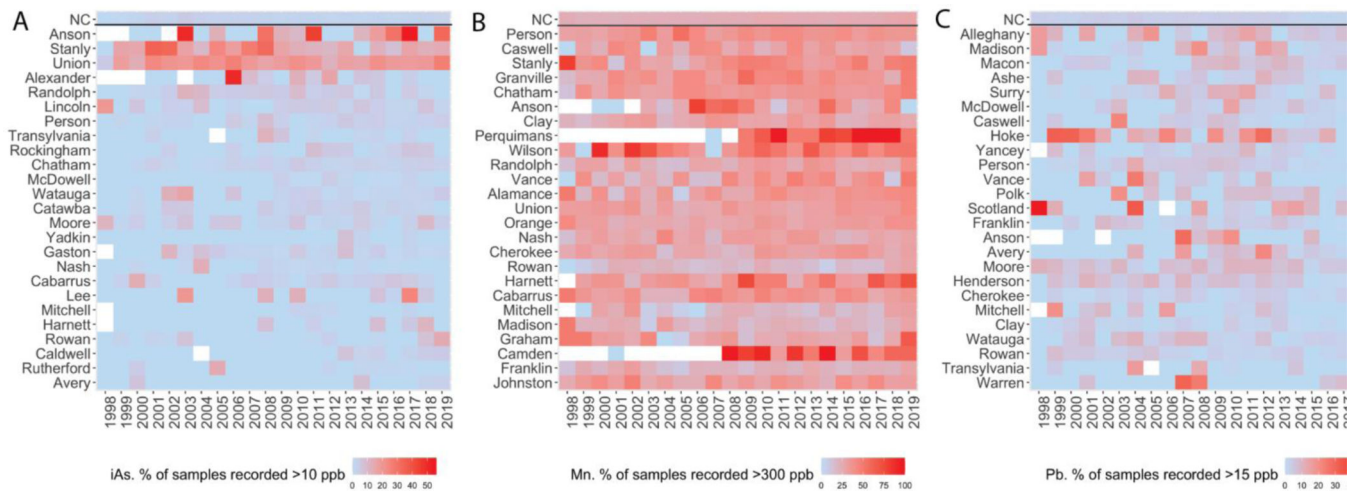


**Figure 3. Bivariate maps of North Carolina demonstrating key correlated pairs of metals in which one metal is toxic.**

Maps show the mean concentrations (in ppb) of (A) arsenic and manganese and (B) lead and copper, by tertiles, summarizing well water tests from October 19, 1998, to May 20, 2019.



**Figure 4. Four clusters of counties representing distinct profiles of county-level mean concentrations for metals with toxic properties were derived.** (A) a map of North Carolina demonstrating the spatial distributions of the clusters, (B) A bar graph demonstrating the mean of the z-score standardized mean concentration of each metal for the counties within each of the three clusters (y=0 represents the state mean) (C) boxplots of the distribution of the percentage of tests over the EPA MCL/SMCL/HAL within the counties in each cluster for each metal.



**Figure 5. The top twenty-five counties for percentage of exceedances of the EPA MCL/HAL for (A) arsenic, (B) manganese and (C) lead.**

Counties are ranked by the percentage of well tests reported that exceed the standard over the entire data collection period. The color-scale, indicated by the legend below each plot, indicate the percentage of exceedances of the MCL (iAs, Pb) or HAL (Mn) for each year. The percentage of exceedances of the EPA MCL/HAL for the entire state for each year are represented by the top row in each plot.



**Table 1.**  
**State-wide distributions of metals in North Carolina private wells that have either a state or federal standard.**

The number of test reports in the dataset, the number and percentage of tests above limit of reporting (LOR), number and percentage of tests at or above standards and maximum reported concentration are provided. Well water tests from October 19, 1998, to May 20, 2019, are included. The number of tests at or above the federal EPA standards or the NC standards are detailed. EPA standards include EPA Primary (MCL) or Secondary Drinking Water standards (SMCL), the Radionuclides rule, EPA Health Advisory Levels, and the EPA Action Level for Lead. NC standards include the NC Groundwater standards, NC Coal Ash Management Act of 2014 (CAMA) Rule screening levels and NC DHHS Health Advisory Level. Only metals with over 200 tests are included.

Metal	Number of tests (N)	Above LOR, N (%)	Maximum (ppb)	EPA standards		NC Groundwater Standards	
				Standard (ppb)	At or Exceed Standard, N (%)	Standard (ppb)	At or Exceed Standard, N (%)
Aluminum	265	81 (30.6)	9090	200 <sup>1</sup>	153 (57.7)	3500 <sup>4</sup>	4 (1.5)
Antimony	502	3 (0.6)	3.54	6 <sup>2</sup>	0 (0)	1 <sup>4,5</sup>	485 (96.4)
Arsenic	110383	7199 (6.5)	806	10 <sup>2,3,4</sup>	2502 (2.3)	10 <sup>2,3,4</sup>	2502 (2.3)
Barium	65439	5022 (7.7)	3400	2000 <sup>2</sup>	7 (<0.1)	2000 <sup>2</sup>	7 (<0.1)
Beryllium	493	0 (0)	2.12	4 <sup>2,4,5</sup>	0 (0)	4 <sup>2,4,5</sup>	0 (0)
Cadmium	75393	255 (0.3)	225	5 <sup>2</sup>	83 (0.1)	2 <sup>3,4</sup>	2148 (2.9)
Chromium	65451	996 (1.5)	760	100 <sup>2</sup>	47 (0.1)	10 <sup>3,4</sup>	930 (1.4)
Copper	65754	7277 (11.0)	277000	1300 <sup>2</sup>	248 (0.4)	1000 <sup>3,4</sup>	342 (0.5)
Iron	65654	29400 (44.8)	221000000	300 <sup>1,3,4</sup>	18732 (28.5)	300 <sup>1,3,4</sup>	18732 (28.5)
Lead	110409	8892 (8.1)	105440	15 <sup>9,3,4</sup>	3076 (2.8)	15 <sup>2,3,4</sup>	3076 (2.8)
Manganese	110182	33653 (30.5)	46300	50 <sup>1,3,4</sup>	26816 (24.3)	50 <sup>1,3,4</sup>	26816 (24.3)
				300 <sup>6</sup>	5405 (4.9)	300 <sup>6</sup>	5405 (4.9)
Mercury	55146	109 (0.2)	5.7	2 <sup>2</sup>	19 (<0.1)	1 <sup>3,4</sup>	54 (0.1)
Selenium	65435	812 (1.2)	623	50 <sup>2</sup>	29 (<0.1)	20 <sup>3,4</sup>	119 (0.2)
Silver	65212	2 (<0.1)	50	100 <sup>1</sup>	0 (0)	20 <sup>3</sup>	65212 (100.0)
Thallium	501	2 (0.4)	1.41	2 <sup>2</sup>	0 (0)	0.2 <sup>4,5</sup>	281 (56.1)
Uranium	580	280 (48.3)	1970	30 <sup>7</sup>	52 (9.0)	30 <sup>7</sup>	52 (9.0)
Zinc	65277	21973 (33.7)	141000000	5000 <sup>1</sup>	1194 (1.8)	1000 <sup>3</sup>	4481 (6.9)
Boron	224	31		~	~	700 <sup>3,4</sup>	0 (0)
		(13.8)					

Metal	Number of tests (N)	Above LOR, N (%)	Maximum (ppb)	EPA standards		NC Groundwater Standards	
				Standard (ppb)	At or Exceed Standard, N (%)	Standard (ppb)	At or Exceed Standard, N (%)
Cobalt	241	20 (8.3)	7.07	~	~	<sup>1</sup> 3.4	38 (15.8)
Hexavalent Chromium	271	166 (61.3)	12.26	~	~	0.07 <sup>8</sup>	148 (54.6)
Molybdenum	223	13 (5.8)	31	~	~	18 <sup>4</sup>	3 (1.4)
Nickel	265	17 (6.4)	50	~	~	100 <sup>3,4</sup>	0 (0)
Strontium	227	47 (20.7)	720	~	~	2100 <sup>4</sup>	0 (0)
Vanadium	237	162 (68.4)	40	~	~	7 <sup>3</sup>	75 (31.7)
				~	~	0.3 <sup>4</sup>	190 (80.2)

~ No EPA standard exists

<sup>1</sup> EPA Secondary Drinking Water Standards that are non-enforceable

<sup>2</sup> EPA Primary Drinking Water Standards that are enforceable

<sup>3</sup> NC DEQ Groundwater Standard

<sup>4</sup> NC CAMA Rule Screening Level

<sup>5</sup> NC IMAC Standard

<sup>6</sup> EPA Health Advisory Level (HAL)

<sup>7</sup> EPA Radionuclides Rule

<sup>8</sup> NC DHHS Health Advisory Level (HAL)

<sup>9</sup> EPA Action Level

**Table 2.**  
**County-level distributions for priority metals, (A) arsenic, (B) manganese, and (C) lead.**

Mean, max, number and percentage over LOR, number and percentage over EPA standard and population in county utilizing private wells are listed for the 25 counties with the highest population-at-risk rank. Population-at-risk rank was determined by multiplying the percentage of wells exceeding the EPA standard and the percentage of population within a county using private wells. PAR= population-at-risk. EOS= Exceedance of standard

County	Number of tests (N)	Above LOR, N (%)	At or Exceed Standard, N (%)	Population using private wells, in thousands (% of county population)	Mean (ppb)	Maximum (ppb)	PAR rank	EOS rank
<b>Arsenic</b>								
<i>State-wide</i>	110383	7199 (6.5)	2502 (2.3)	-	3.4	806	-	-
Anson	132	40 (30.3)	25 (18.9)	8.6 (34.6)	10.9	154	1	2
Stanly	1278	482 (37.7)	230 (18.0)	19.3 (31.2)	8.9	806	2	3
Union	5154	1703 (33.0)	1003 (19.5)	49.2 (20.9)	8.7	303	3	1
Alexander	272	39 (14.3)	16 (5.9)	13.3 (35.6)	5.0	175	4	5
Randolph	2619	379 (14.5)	126 (4.8)	60.1 (41.9)	3.8	106	5	6
Lincoln	2333	174 (7.5)	85 (3.6)	43.5 (52.0)	4.3	390	6	9
Person	1804	156 (8.7)	45 (2.5)	27.1 (68.6)	3.4	54	7	20
Transylvania	580	25 (4.3)	14 (2.4)	17.0 (49.8)	3.5	68	8	22
Rockingham	986	76 (7.7)	25 (2.5)	38.8 (42.8)	3.3	47	9	19
Chatham	3129	422 (13.5)	85 (2.7)	28.9 (40.0)	3.6	127	10	17
McDowell	1251	40 (3.2)	16 (1.3)	31.1 (68.3)	3.5	93	11	31
Watauga	1077	68 (6.31)	26 (2.4)	18.8 (33.5)	3.3	38	12	23
Catawba	1567	87 (5.6)	44 (2.8)	45.1 (28.5)	4.6	388	13	15
Moore	1280	133 (10.6)	41 (3.2)	24.6 (24.9)	3.8	167	14	11
Yadkin	328	7 (2.1)	3 (0.9)	25.7 (68.6)	3.2	25	15	40
Gaston	2764	199 (7.2)	77 (2.8)	49.1 (22.0)	3.4	232	16	16
Nash	2258	115 (5.1)	35 (1.6)	34.7 (36.9)	3.1	32	17	28
Cabarrus	1382	126 (9.1)	44 (3.2)	37.9 (17.9)	3.6	247	18	12
Lee	278	54 (19.4)	10 (3.6)	9.4 (15.3)	3.5	29	19	10
Mitchell	564	13 (2.3)	5 (0.9)	9.1 (60.7)	3.2	39	20	41
Harnett	368	25 (6.8)	9 (2.5)	29.1 (21.7)	3.6	127	21	21
Rowan	2374	108 (4.6)	24 (1.0)	69.9 (49.5)	2.7	52	22	38
Caldwell	358	15 (4.2)	8 (2.2)	16.4 (19.9)	3.6	27	23	24
Rutherford	784	19 (2.4)	8 (1.0)	27.4 (41.1)	3.3	46	24	37
Avery	640	19 (3.0)	7 (1.1)	6.3 (36.1)	4.7	484	25	34
<b>Manganese</b>								
<i>State-wide</i>	110182	33653 (6.5)	5405 (4.9)	-	78.1	46300	-	-
Person	1803	797 (44.2)	255 (14.1)	27.1 (68.6)	147.0	11000	1	4

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

County	Number of tests (N)	Above LOR, N (%)	At or Exceed Standard, N (%)	Population using private wells, in thousands (% of county population)	Mean (ppb)	Maximum (ppb)	PAR rank	EOS rank
Caswell	688	250 (36.3)	45 (6.5)	19.4 (85.4)	146.0	44950	2	20
Stanly	1230	478 (38.9)	211 (17.2)	19.3 (31.2)	177.0	12180	3	2
Granville	1222	548 (44.8)	137 (11.2)	28.4 (47.3)	137.7	10000	4	7
Chatham	3128	1320 (42.2)	321 (10.3)	28.9 (39.5)	120.6	3900	5	12
Anson	136	55 (40.4)	15 (11.0)	8.6 (34.6)	128.7	4140	6	9
Clay	1575	547 (34.7)	75 (4.8)	8.8 (79.1)	84.1	5850	7	30
Perquimans	52	45 (86.5)	11 (21.2)	2.3 (16.9)	210.0	1470	8	1
Wilson	465	263 (56.6)	70 (15.1)	18.5 (22.7)	155.6	2000	9	3
Randolph	2609	767 (29.4)	210 (8.2)	60.2 (41.9)	117.6	11000	10	15
Vance	387	142 (36.7)	22 (5.7)	25.2 (56.6)	81.6	1900	11	26
Alamance	2600	1072 (41.2)	249 (9.6)	49.9 (30.0)	109.2	4320	12	14
Union	5137	1763 (34.3)	679 (13.2)	49.2 (20.9)	147.3	6730	13	5
Orange	7250	2485 (34.3)	763 (10.5)	37.8 (25.9)	111.0	10300	14	10
Nash	2260	787 (34.8)	152 (6.7)	34.7 (36.9)	93.4	10670	15	19
Cherokee	3031	1034 (34.1)	116 (3.8)	17.6 (62.2)	71.2	5800	16	35
Rowan	2379	488 (20.5)	105 (4.4)	69.9 (49.5)	73.8	6000	17	32
Harnett	368	182 (49.5)	36 (9.8)	29.1 (21.7)	106.8	4400	18	13
Cabarrus	1399	582 (41.6)	147 (10.5)	37.9 (17.9)	127.2	6190	19	11
Mitchell	564	165 (29.3)	17 (3.0)	9.1 (60.7)	134.5	11000	20	38
Madison	1190	313 (26.3)	33 (2.8)	14.3 (65.7)	51.7	1800	21	43
Graham	454	145 (31.9)	16 (3.5)	4.2 (49.5)	88.6	6140	22	37
Camden	92	59 (64.1)	6 (6.5)	2.8 (26.3)	116.8	490	23	21
Franklin	1559	424 (27.2)	41 (2.6)	42.6 (63.0)	55.9	2560	24	44
Johnston	2002	754 (37.7)	145 (7.2)	46.3 (22.8)	89.3	3170	25	16
<b>Lead</b>								
<i>State-wide</i>	<i>110409</i>	<i>8892 (8.1)</i>	<i>3076 (2.8)</i>	<i>-</i>	<i>7.3</i>	<i>105440</i>	<i>-</i>	<i>-</i>
Alleghany	347	84 (24.2)	24 (6.9)	7.8 (69.6)	8.7	707	1	5
Madison	1190	165 (13.9)	76 (6.4)	14.3 (65.7)	7.7	701	2	8
Macon	2424	345 (14.2)	133 (5.5)	20.9 (59.3)	8.9	2540	3	10
Ashe	717	72 (10.0)	26 (3.6)	22.3 (82.1)	8.0	2300	4	23
Surry	1171	138 (11.8)	49 (4.2)	47.1 (65.5)	5.6	263	5	18
McDowell	1250	113 (9.0)	49 (3.9)	31.1 (68.3)	6.8	974	6	21
Caswell	688	75 (10.9%)	20 (2.9)	19.4 (85.4)	4.7	80	7	33
Hoke	306	76 (24.8)	39 (12.8)	9.7 (17.7)	9.1	191	8	2
Yancey	1029	105 (10.2)	32 (3.1)	12.4 (69.2)	4.9	259	9	28
Person	1802	164 (9.1)	56 (3.1)	27.1 (68.6)	5.3	304	10	27
Vance	387	40 (10.3)	13 (3.4)	25.2 (56.6)	4.7	57	11	25

County	Number of tests (N)	Above LOR, N (%)	At or Exceed Standard, N (%)	Population using private wells, in thousands (% of county population)	Mean (ppb)	Maximum (ppb)	PAR rank	EOS rank
Polk	401	38 (9.5)	16 (4.0)	9.7 (47.3)	5.6	162	12	19
Scotland	266	67 (25.2)	16 (6.0)	10.9 (31.2)	9.5	761	13	9
Franklin	1560	185 (11.9)	46 (3.0)	42.6 (63.0)	5.0	146	14	32
Anson	131	19 (14.5)	7 (5.3)	8.6 (34.6)	6.9	260	15	11
Avery	640	75 (11.7)	30 (4.7)	6.3 (36.1)	5.4	147	16	15
Moore	1289	223 (17.3)	86 (6.7)	24.6 (24.9)	6.3	205	17	7
Henderson	1497	201 (13.4)	71 (4.7)	40.8 (35.0)	6.1	355	18	14
Cherokee	3035	213 (7.0)	76 (2.5)	17.6 (62.2)	4.5	222	19	47
Mitchell	564	55 (9.8)	14 (2.5)	9.1 (60.7)	4.4	68	20	49
Clay	1575	109 (6.9)	30 (1.9)	8.8 (79.1)	4.5	216	21	68
Watauga	1065	113 (10.6)	41 (3.9)	18.8 (33.5)	8.2	2660	22	22
Rowan	2378	175 (7.4)	62 (2.6)	69.9 (49.5)	5.1	587	23	42
Transylvania	580	38 (6.6)	15 (2.6)	17.0 (49.8)	6.1	397	24	43
Warren	292	30 (10.3)	9 (3.1)	8.2 (41.3)	5.2	161	25	29

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3.**  
**One-way ANOVA results for testing differences in metal concentrations between four clusters.**

Post-hoc pairwise t-tests to compare individual cluster pairs were also conducted. P-values are adjusted using the Benjamini-Hochburg procedure for conducting ANOVA and subsequent post-hoc pairwise t-tests on six different metals.

	ANOVA test of difference in means between all four clusters		Post-hoc pairwise t-test comparison of means between individual cluster pairs					
			1 vs 2	2 vs 3	1 vs 3	1 vs 4	2 vs 4	3 vs 4
Metal	F statistic	BH-adjusted p value	BH-adjusted p value	BH-adjusted p value	BH-adjusted p value	BH-adjusted p value	BH-adjusted p value	BH-adjusted p value
Arsenic	127.49	1.39E-32	1.00E+00	1.74E-01	2.94E-01	2.10E-34	5.43E-33	1.65E-28
Manganese	5.75	6.96E-03	1.00E+00	1.00E+00	1.00E+00	2.20E-04	6.95E-04	1.52E-02
Lead	23.62	9.35E-11	1.00E+00	1.39E-11	1.28E-12	1.00E+00	1.00E+00	7.05E-07
Cadmium	28.58	1.63E-12	1.00E+00	7.75E-14	3.76E-14	1.00E+00	1.00E+00	4.01E-08
Mercury	155.22	6.45E-36	1.30E-37	6.55E-14	1.00E+00	1.00E+00	1.31E-11	1.00E+00
Chromium	1.39	1.00E+00	1.00E+00	1.00E+00	7.01E-01	3.53E-01	7.29E-01	1.00E+00