

RESEARCH ARTICLE

Open Access



# Lead contamination of public drinking water and academic achievements among children in Massachusetts: a panel study

Wenxin Lu, Ronnie Levin\*  and Joel Schwartz

## Abstract

**Background:** Public drinking water can be an important source exposure to lead, which can affect children's cognitive development and academic performance. Few studies have looked at the impact of lead exposures from community water supplies or their impact on school achievements. We examined the association between annual community water lead levels (WLLs) and children's academic performances at the school district level.

**Methods:** We matched the 90th percentile WLLs with the grade 3–8 standardized test scores from the Stanford Education Data Archive on Geographic School Districts by geographic location and year. We used multivariate linear regression and adjusted for urbanicity, race, socioeconomic characteristics, school district, grade, and year. We also explored potential effect measure modifications and lag effects.

**Results:** After adjusting for potential confounders, a 5 µg/L increase in 90th percentile WLLs in a GSD was associated with a 0.00684 [0.00021, 0.01348] standard deviation decrease in the average math test score in the same year. No association was found for English Language Arts.

**Conclusions:** We found an association between the annual fluctuation of WLLs and math test scores in Massachusetts school districts, after adjusting for confounding by urbanicity, race, socioeconomic factors, school district, grade, and year. The implications of a detectable effect of WLLs on academic performance even at the modest levels evident in MA are significant and timely. Persistent efforts should be made to further reduce lead in drinking water.

**Keywords:** Environmental epidemiology, Lead, Water, Children, Education, Cognitive development, Academic performance

## Background

Lead is a prevalent environmental contaminant and a known neurotoxic agent. Lead crosses the blood brain barrier [1], interferes with the calcium-regulated release of neurotransmitters [2] and induces programmed cell death of the nervous system [3]. Developing brains of children are particularly vulnerable to the neurotoxic effects of lead. Exposure to even low levels of lead shows

evidence of long-term damage to children's cognitive function and IQ [4–10], and children's academic performance [11–18], with implications for future academic and career achievements [19, 20].

Despite widespread reported compliance with EPA's 1991 Lead and Copper Rule (LCR), as administered by state drinking water agencies, drinking water remains an important potential lead exposure source [21–23]. Lead rarely occurs in natural water sources, it contaminates drinking water via the corrosion of lead pipes, solder, faucets, cisterns, and other plumbing components containing lead. Exposure to lead from drinking

\*Correspondence: rblevin@hsph.harvard.edu  
Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, USA



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

water has been associated with variabilities in children's blood lead levels (BLLs) [24–27]. Interventions, such as lead pipe replacement, can significantly reduce WLLs [28], and consequently, BLLs [29]. EPA estimated that drinking water generally constitutes more than 20% of average daily lead exposure, 40 to 60% for infants who consume mostly infant formula (dry powder or liquid concentrate) mixed with tap water, and up to 80% of children's daily exposure in some realistic circumstances even in public water supplies (PWSs) that are not exceeding EPA's LCR [30]. Variability, especially in water sample collection methods, complicates comparing studies on the relationship between BLLs and WLLs [22]. Quantifying the contribution of WLLs to BLLs in children is further complicated because of the difficulty of collecting reproducible water lead samples; relatively small fluctuations in factors such as temperature, pH, alkalinity, and dissolved solids affect the solubility of lead [31, 32].

Exposure to other sources of lead, such as the legacy lead deposition in soils, have been found to affect the cognitive abilities of young children [33]. Yet few studies have attempted to directly quantify the health impact of WLLs. We investigated whether water system wide WLLs are reflected in the academic performances of children on a population level, without information on individual blood lead concentrations. We found only one study investigating a short-term water-lead exposure with school performance [34], but that was unpublished.

Most previous studies relating lead exposure with child academic performances focused on long-term cumulative effects. For instance, Miranda et al. found that BLLs in early childhood (0–5 years old) were associated with lower 4th grade math and reading test scores [35]. Reyes showed that children in Massachusetts with higher BLLs in early childhood (before age 6) perform worse in standardized tests in grades 3 and 4 [36]. Aizer et al. showed that reducing BLLs before the age of 6 reduces the probability of low reading and math proficiency at 3rd grade [37]. A New York study found county-level incidence of higher BLLs is associated with lower test scores 3–7 years later in grades 3–8 [13]. To the best of our knowledge, no one has investigated the potential impact of recent water lead exposures on cognitive function and academic performance.

To examine whether recent WLLs are associated with children's test scores on the population level, we matched the community WLLs to the test scores of students from grade 3 to grade 8 in Massachusetts by school district. We constructed a panel dataset using repeated measurements from 2010 to 2016, and tested math and English Language Arts (ELA) separately. We adjusted for urbanicity, race, SES and other potential confounders.

## Methods

### Study Population

The study region is the Commonwealth of Massachusetts, USA, which is partitioned into 294 Geographic School Districts (GSD). GSDs with incomplete information on geographical coverage were removed. We restricted to GSDs that are served by at least one Community Water System (CWS), i.e., a public water system that supplies water year-round to at least 25 people at their primary residence or at least 15 primary residences [38]. Standardized test score data and lead concentration data were available for the remaining 229 GSDs from 2009 to 2016. Because GSD-level covariates data are generally missing for year 2009, we restricted to the 7-year period from 2010 to 2016, constituting 2,164,208 student-years during the period. To test whether restricting to the 229 GSDs resulted in any selection bias, we compared the average test score among the 229 included GSDs and the 65 excluded GSDs. The average standardized test scores pooled across subjects are 0.482 and 0.479 for the included and excluded GSDs, respectively. A two-sample t-test showed no significant difference between the two groups [ $p = 0.95$ ]. Since the outcome is uncorrelated with whether the GSD was included in the analysis, the possibility of selection bias is minimal.

### Lead concentration data

The 90th percentile lead concentration data in each CWS was obtained from the MA Department of Environmental Protection (DEP), which is the primacy agency for the state of MA under the Safe Drinking Water Act. The federal LCR requires, among other conditions, that WLLs in a CWS not exceed the Action Level (AL) of 15  $\mu\text{g}/\text{L}$  at the 90th percentile across all samples in a monitoring period [30].

Under the LCR, each CWS must submit two documents for DEP approval before sampling begins: A Materials Evaluation that surveys the material and structure characteristics of a pool of target sampling sites; and a Sampling Plan that identifies the sampling sites at the highest risks of elevated lead concentrations based on their structure characteristics, including the likelihood of lead service lines, solder, and goosenecks. The majority of sampling sites for CWSs are households, including single-family and multi-family residences. Under the federal LCR, the water sample from each residence in the sample pool is a one liter first-flush sample taken after a minimum 6-h stagnation interval to approximate the routine high exposures that can occur in each home every morning and/or at the end of the day when residents return home from work or school. Lead analysis is conducted at certified laboratories. The sample lead concentrations are reported to the DEP and the 90th percentile lead

concentrations for each CWS are recorded. During each monitoring period, the CWS must collect a prescribed number of 1-1 first-draw water samples from their available sampling sites. The number of sampling sites required depends on the size of the population served (Table 1) [39].

Also, under the LCR, there are 3 monitoring periods annually: Q2 (Jan 1st to Jun 30th), Q3 (Jun 1st to Sep 30th) and Q4 (Jul 1st to Dec 31st). The standard testing schedule for CWSs is semiannual testing, which happens in Q2 and Q4 each year. If a CWS has met both the lead and copper requirements for two consecutive semi-annual monitoring periods, monitoring frequency is reduced from semi-annual to annual. If a CWS has met both the lead and copper requirements for three consecutive annual monitoring periods, its monitoring frequency can be further reduced from annual to triennial. If the lead or copper requirement is violated during a reduced-frequency monitoring period, the monitoring frequency is reversed to semiannual. An increase in monitoring frequencies can also result from changes in source water, treatment methods and overall operations of the water systems, or failing to operate within their water quality parameters for more than nine days in any six-month period. When the LCR is violated, the CWSs are required to undergo interventions such as corrosion control treatment, lead service line replacement and public education. Systems under annual and triennial testing sample during the monitoring period Q3, and the number of sampling sites required in these systems are also reduced, as described in Table 1. The requirement that reduced sampling should be implemented during summer months was proposed in the Preamble to the 1991 LCR, because plumbosolvency is increased by temperature, resulting in higher WLLs in the summer [32, 40].

The MA Water Resources Authority (MWRA) is the largest water supplier in MA, serving approximately 2,222,151 people in 2010 [41]. For CWSs served by the MWRA, the Q3 monitoring period is from July 1st

to October 31st, which is slightly different from the Q3 monitoring period of other CWSs independent of MWRA. Of the 351 cities and towns in Massachusetts, 32 cities and towns, including the most populous city Boston, are fully supplied by MWRA; another 15 cities and towns are partially supplied by MWRA [42].

The DEP data cover 438 active CWSs with 1847 drinking water sources, serving a total of approximately 4,819,215 people. As a result of the different sampling frequencies, the monitoring periods in the data are of varying lengths. Among all MA CWSs from 2010 to 2016, there were 408 semi-annual monitoring periods, 446 annual monitoring periods, 1180 triennial monitoring periods and 10 monitoring periods of other lengths. About 95.1% of the data were collected in the Q3 (summer) monitoring period. To address monitoring periods of varying lengths, we calculated the 90th percentile WLL for each CWS-year as a weighted average of the 90th percentile lead concentrations of all monitoring periods in that year, weighted on the number of days in the monitoring periods covered in that year. For this study, we assumed the 90th percentile WLLs are indicative of the relative ranking of community exposure. Similar WLL data and assumptions have been used in other published studies [43, 44]. For simplicity, we will refer to the 90th percentile WLL as the WLL in this analysis. The WLL for each GSD-year was calculated as the average of the WLLs of the active CWSs located within the GSD, weighted on the population served by each CWS site. The geographic location of the CWS sites was retrieved from the MA Bureau of Geographic Information (MASSGIS) [45]. Data on population served by each CWS was retrieved from the Environmental Working Group’s Tap Water Database [46].

**Test score and covariates**

The GSD level standardized test score data was retrieved from the Stanford Education Data Archive (SEDA) version 3.0 [47]. Individual-level data were not available

**Table 1** Number of required sampling sites required for community water system sizes, for standard monitoring (semi-annual) and reduced monitoring (annual and triennial)

Size of CWS (People served)	Number of sampling sites (standard monitoring)	Number of sampling sites (reduced monitoring)
> 100,000	100	50
10,001 – 100,000	60	30
3301 – 10,000	40	20
501–3300	20	10
101–500	10	5
≤ 100	5	5

for the protection of the privacy of student educational records, required by The Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99) [48]. We chose data aggregated on the GSD-level because it was the highest resolution with annual data available. The data we obtained contains GSD-level academic achievement information measured as standardized test scores in Mathematics and ELA, which in MA is the Massachusetts Comprehensive Assessment System (MCAS). The scale of the student test scores can be interpreted as the number of standard deviations (SD) above the average student performance, compared to a reference cohort (students in the 4th grade in 2009). For instance, a GSD-grade with a 0.5 average standardized test score indicates that the students of that grade in that GSD performed on average 0.5 SD higher than the average test score of the reference cohort in the same grade. The standardized test score for all GSDs in the US follows an approximate normal distribution. In this study, we included only GSDs in MA. The median MA standardized test score is approximately 0.49 for math and 0.47 for ELA, indicating that MA students perform above the national average by grade. Data are available for each grade-year-subject from grade 3 to grade 8 and from 2008 to 09 to 2015–16 school years. After restricting to the 229 GSDs with complete geographic information and served by at least one CWS, the remaining dataset covers 2,164,028 enrolled student-years and 5,677,721 tests. For each enrolled student-year, two tests are administered (math test and ELA test), thus the number of tests is approximately twice the number of student-years. The actual ratio of tests versus student-years is slightly higher than 2, due to practical arrangements such as retests.

Covariates data with the SEDA datasets are from the Common Core of Data of the National Center for Education Statistics and the American Community Survey. The covariates included in the models in this study include four categories: (1) urbanicity: urban, sub-urban, town and rural; (2) GSD SES characteristics: log of median income, bachelor's degree rate, poverty rate, Supplemental Nutrition Assistance Program (SNAP) recipient rate, single-mother household rate and unemployment rate; (3) student racial composition: proportion of native Americans, Asians, Hispanics, Blacks and Whites; and (4) student SES characteristics: proportion of English Language Learners (ELL), proportion of reduced-price lunch eligible students and proportion of economically disadvantaged students.

### Statistical analysis

We used a multivariate linear regression with fixed effects for time and cohort, which is a variant of the causal Difference-in-Difference (DID) analysis [49–51].

In this study, a cohort is defined as all students enrolled in a specific grade in a specific GSD, and time is school year. This model controls for unmeasured confounders across GSDs, grades within GSDs, and time periods by adding dummy variables for each cohort and year [52]. Fixed or slowly varying covariates within GSD are also removed by the cohort dummy variables. The exposure of interest is WLL for each cohort within each year; the outcome of interest is cohort-year standardized test scores. Thus, after controlling for cohort fixed effects and year fixed effects, we tested whether deviations in WLL from the cohort-year average are associated with deviations in standardized test scores from the cohort-year average.

The base model takes the following form:

$$S_{itg} = \beta_1 WLL_{it} + \beta_2 Z_{itg} + \alpha_{ig} + \gamma_t + \epsilon_{itg}$$

where  $i$  indexes GSD,  $t$  indexes school year and  $g$  indexes grade. The dependent variable  $S_{itg}$  is the standardized test score for grade  $g$  students in GSD  $i$  in year  $t$ . The independent variable of interest  $WLL_{it}$  is the average WLL in GSD  $i$  in year  $t$ .  $Z_{itg}$  are cohort-level covariates associated with academic performances for grade  $g$  students in GSD  $i$  in year  $t$ .  $\alpha_{ig}$  are cohort fixed effects for GSD  $i$  and grade  $g$ , and  $\gamma_t$  are year fixed effects for year  $t$ . Because  $Z_{itg}$  are time varying, adjusting for these covariates controls for differences across GSDs due to differential changes of these covariates over time. We used linear models in panel data in R version 4.0.5 (package “plm”).

This model is generalized from a traditional DID model by extending the scope to more than two time periods and investigating a continuous treatment variable. The continuous treatment is the change in WLL for a GSD over two consecutive years, and the outcome is the change in standardized test scores for a GSD over the same two years. Consider two cohorts A and B over two consecutive time periods  $t$  and  $t+1$ . Assume the exposure level of cohort A does not change from  $t$  to  $t+1$ , while the exposure level of cohort B increased by 1 unit from  $t$  to  $t+1$ . In this scenario, cohort A can be considered a negative outcome control for cohort B: changes in the outcomes in cohort A cannot be due to variations in the exposure levels, and can only be explained by variations in the same set of covariates as cohort B. For both cohorts A and B, the outcome levels over the two time periods, the expected differences, and DID estimates are presented in Table 2. The linear DID estimate between the two cohorts over the two time periods is:

$$\beta_1 + \beta_2 [(Z_{B,t+1} - Z_{Bt}) - (Z_{A,t+1} - Z_{At})]$$

Under the assumption that the time varying covariates  $Z_{itg}$  capture all other factors affecting the trend of the outcome and that there are no interactions between



**Table 2** linear DID estimation process with two-way fixed effects models

	Cohort A (no change in WLL)	Cohort B (WLL increased by 1 unit)
t	$\beta_1 WLL_{At} + \beta_2 Z_{At} + \alpha_A + \gamma_t + \epsilon_{At}$	$\beta_1 WLL_{Bt} + \beta_2 Z_{Bt} + \alpha_B + \gamma_t + \epsilon_{Bt}$
t + 1	$\beta_1 WLL_{A,t+1} + \beta_2 Z_{A,t+1} + \alpha_A + \gamma_{t+1} + \epsilon_{A,t+1}$	$\beta_1 (WLL_{Bt} + 1) + \beta_2 Z_{B,t+1} + \alpha_B + \gamma_{t+1} + \epsilon_{B,t+1}$
E(Difference)	$\beta_2 (Z_{A,t+1} - Z_{At}) + (\gamma_{t+1} - \gamma_t)$	$\beta_1 + \beta_2 (Z_{B,t+1} - Z_{Bt}) + (\gamma_{t+1} - \gamma_t)$
DID	$\beta_1 + \beta_2 [(Z_{B,t+1} - Z_{Bt}) - (Z_{A,t+1} - Z_{At})]$	

cohort fixed effects and year fixed effects [53], the estimated coefficient  $\beta_1$  is an estimate of the causal treatment effect of a 1-unit increase of WLL over the time periods t to t + 1. This method was used in a recent paper investigating the association between ambient air pollution and children’s academic performances in the US [54].

To test the potential lag effects of WLLs, we replaced the WLLs in the model with the WLL level from the previous year. The model takes the following form:

$$S_{itg} = \beta_1 WLL_{i,t-1} + \beta_2 Z_{itg} + \alpha_{ig} + \gamma_t + \epsilon_{itg}$$

Where  $WLL_{i,t-1}$  is the average WLL in GSD i in year t-1. The coefficient  $\beta_1$  and its statistical significance is compared with the respective coefficient from the results of the base model.

To test potential effect measure modifications by grade and racial composition, we modified the base model by adding an interaction term between the WLLs and the potential effect modifier. The model takes the following form:

$$S_{itg} = \beta_1 WLL_{it} + \beta_2 WLL_{it} Y_{itg} + \beta_3 Z_{itg} + \alpha_{ig} + \gamma_t + \epsilon_{itg}$$

Where  $Y_{itg}$  is the level of the potential effect modifier of grade g students in GSD i in year t. The coefficient  $\beta_2$  is tested for statistical significance. A deviation of  $\beta_2$  from zero implies that the association between WLL and test scores is different among observations with different levels of the effect modifier  $Y_{itg}$ .

Because the number of children in each grade in each school districts varies substantially, we used weighted least squares (WLS), where the weight is the number of students enrolled in each cohort within each year. We truncated the weights at the 5th and 95th percentiles to avoid influential observations and to stabilize the weighting. The weights for cohorts with number of enrolled students more than the 95th percentile of all cohorts were replaced by the 95th percentile, which is 596 students; the weights for cohorts with number of enrolled students less than the 5th percentile of all cohorts were replaced by the 5th percentile, which is 25 students.

## Results

### Summary Statistics

Table 3 presents the summary statistics of standardized test scores, number of student enrollments, average WLLs and socioeconomic characteristics of the 229 GSDs in Massachusetts, averaged across grades 3–8 and years 2010–2016. Urbanicity and GSD SES characteristics vary by district and year, while student racial composition and student SES characteristics vary by district, year and grade. The median standardized math and ELA test scores in Massachusetts are 0.4949 and 0.4672 SD above the national median level, respectively.

Table 4 presents the distribution of the standardized test scores (math and ELA) averaged across grades and the average WLL for the 229 GSDs in MA, from 2010 to 2016. The minimum, 10th percentile, median, 90th percentile, maximum and SD for each variable in each year are shown. The median WLLs have a declining trend in general and the maximum WLLs have a slight increasing trend except for year 2013.

Figure 1 presents the box plot distribution of GSD-level WLLs from 2010 to 2016. Each point represents the WLL level for a GSD-year observation. In each calendar year, some GSDs had WLLs above the 15 µg/L action level. In all years except for year 2013, there was one GSD with WLL over 45 µg/L, much higher than all other GSDs in MA.

Figure 2 illustrates the GSD-level standardized test scores for both subjects, pooled over 7 years from 2010 to 2016 and 6 grades from grade 3 to grade 8.

Geographic distribution of GSD-level standardized test score was plotted using ArcGIS Pro 2.8 licensed by Harvard University Center for Geographic Analysis. The shapefiles for GSD maps were retrieved from Stanford Education Data Archive (SEDA) version 3.0 [47].

### Regression Results

We investigated two outcomes of interest: the standardized math test scores and the standardized ELA test scores. The main regression results are presented in

**Table 3** GSD-level summary statistics, averaged across grades and years

	Min	25th percentile	Median	Mean	75th percentile	Max	SD	n
Standardized Math Test Score <sup>a</sup>	−0.417	0.314	0.495	0.495	0.696	1.266	0.299	228
Standardized ELA Test Score <sup>a</sup>	−0.555	0.284	0.467	0.467	0.654	1.170	0.293	228
Number of enrolled students	10.21	96.00	181.19	229.43	289.24	1911	224.45	229
WLLs (µg/L)	0.143	2.349	3.960	4.433	5.793	27.279	3.172	222
Percentage Native American (%)	0.000	0.047	0.136	0.256	0.265	5.568	0.513	229
Percentage Asian (%)	0.000	1.122	2.064	2.064	4.232	29.321	4.346	229
Percentage Hispanic (%)	0.260	2.377	3.792	7.092	6.435	91.086	11.018	229
Percentage Black (%)	0.000	0.902	1.679	0.310	3.113	55.220	5.341	229
Percentage White (%)	5.532	82.748	91.016	85.910	94.653	98.959	14.881	229
Percentage English Language Learner (%)	0.000	0.410	0.994	2.550	2.338	30.016	4.411	228
Percentage Reduced-price Lunch Eligible (%)	0.775	2.193	3.743	4.220	6.137	9.905	2.338	229
Percentage Economically Disadvantaged (%) <sup>b</sup>	1.847	11.803	19.300	23.912	31.740	85.957	16.793	229
Log of Median Income	10.43	11.06	11.25	11.26	11.44	12.05	0.30	227
Bachelor's Degree Rate (%)	11.40	29.32	38.23	40.54	49.42	80.67	15.44	227
Poverty Rate (%)	0.636	3.930	6.824	8.397	11.506	37.943	6.209	227
SNAP Receipt Rate (%) <sup>b</sup>	0.899	3.870	6.018	7.023	8.142	35.674	5.224	227
Single-mom Household Rate (%)	4.52	11.29	13.79	14.36	16.23	43.34	5.28	227
Unemployment Rate (%)	3.424	6.037	6.877	7.109	8.103	14.177	1.745	227

<sup>a</sup> The test scores are standardized on the national level, thus the standardized scores for Massachusetts do not center at 0

<sup>b</sup> Economically disadvantaged students in Massachusetts are defined as students who are participating in one or more of the four state-administered programs: the Supplemental Nutrition Assistance Program (SNAP), the Transitional Assistance for Families with Dependent Children (TAFDC), the Department of Children and Families' (DCF) foster care program, and MassHealth (Medicaid)

**Table 4** Distribution of standardized test scores and WLLs across GSDs, 2010–2016, averaged across grades

		2010	2011	2012	2013	2014	2015	2016
Standardized Math Test Score	Min	−0.488	−0.348	−0.361	−0.346	−0.390	−0.512	−0.529
	10th percentile	0.170	0.176	0.168	0.141	0.091	0.049	0.007
	Median	0.505	0.528	0.537	0.545	0.498	0.420	0.411
	Mean	0.508	0.521	0.540	0.541	0.499	0.441	0.430
	90th percentile	0.879	0.888	0.931	0.958	0.922	0.871	0.897
	Max	1.163	1.199	1.360	1.388	1.324	1.246	1.347
	SD	0.278	0.279	0.303	0.318	0.324	0.316	0.345
Standardized ELA Test Score	Min	−0.642	−0.529	−0.558	−0.588	−0.582	−0.535	−0.449
	10th percentile	0.089	0.125	0.102	0.076	0.093	0.083	0.084
	Median	0.450	0.482	0.481	0.466	0.486	0.479	0.454
	Mean	0.452	0.476	0.476	0.464	0.469	0.467	0.472
	90th percentile	0.854	0.891	0.889	0.903	0.896	0.826	0.896
	Max	1.194	1.115	1.287	1.272	1.227	1.146	1.230
	SD	0.296	0.298	0.307	0.318	0.312	0.303	0.306
90th Percentile WLLs (µg/L)	Min	0	0	0	0	0	0	0
	10th percentile	1.197	1.000	1.278	1.000	1.195	1.000	1.101
	Median	4.000	3.903	3.819	3.122	3.308	3.018	3.300
	Mean	4.760	4.417	4.590	4.203	4.393	4.315	4.452
	90th percentile	8.635	8.000	8.833	8.010	7.859	8.358	7.983
	Max	47.538	46.668	47.830	17.000	50.000	50.000	50.000
	SD	4.376	4.038	4.249	3.044	4.304	4.453	4.681

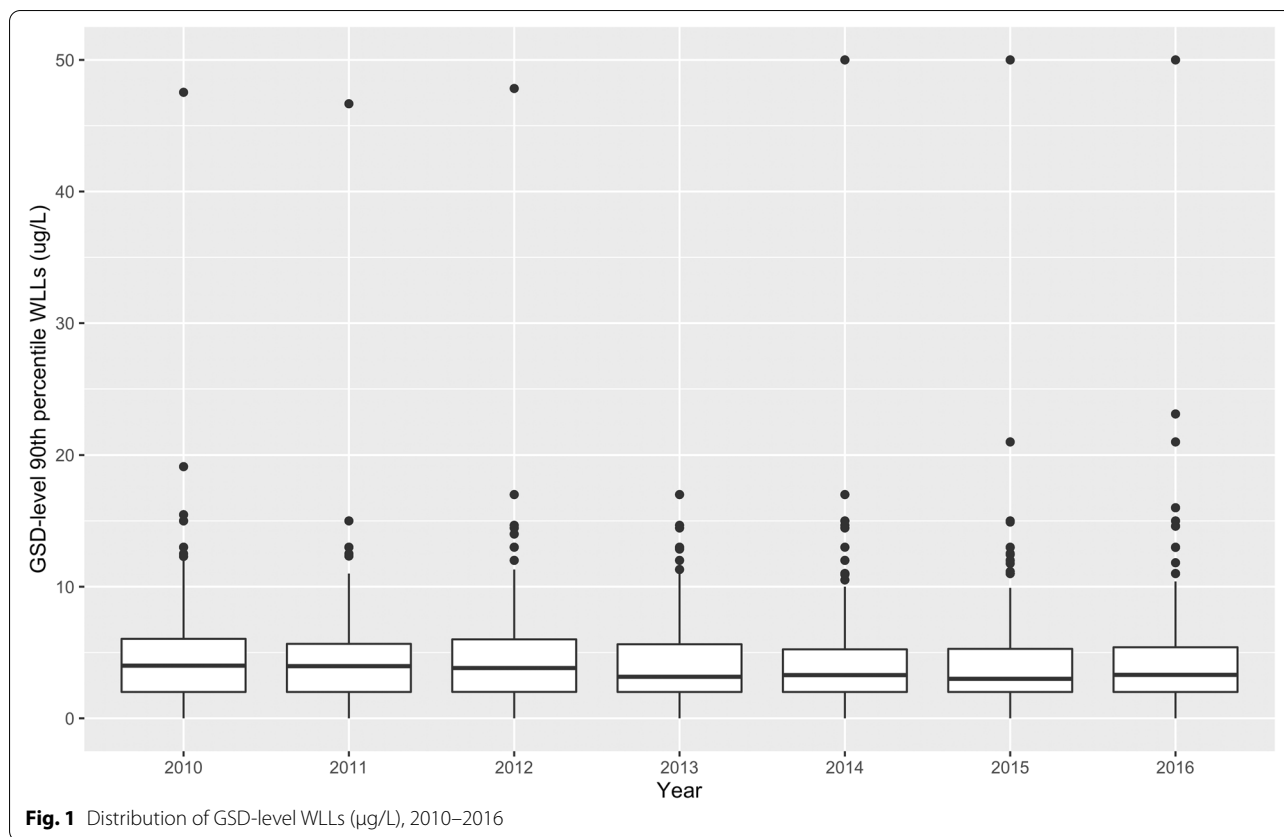
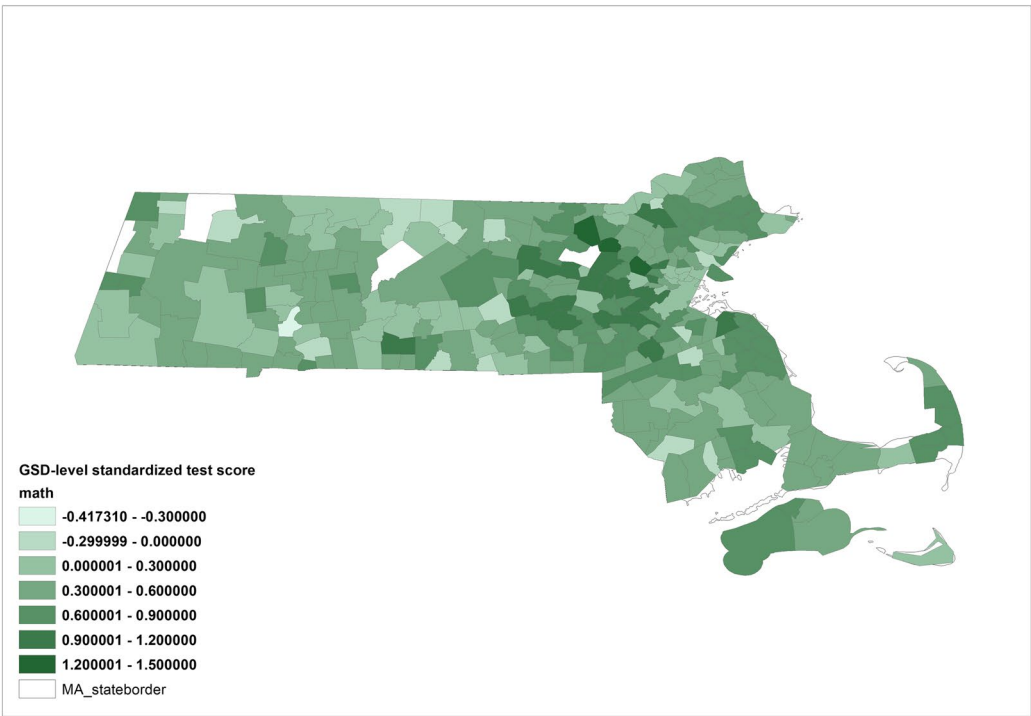


Table 5. For each outcome of interest, three models are presented. The first one is a crude model with only the exposure of interest, the cohort fixed effects and the time fixed effects. Urbanicity, GSD-level SES characteristics, student racial compositions and student SES characteristics are progressively added to the second and third models. For ease of interpretation, results are presented for a 5  $\mu\text{g/L}$  increase in WLL. After adjusting for all the above-mentioned covariates, the cohort fixed effects and year fixed effects (column 3), a 5  $\mu\text{g/L}$  increase in WLL in a GSD is associated with a 0.00684 SD decrease in the standardized math test score in the GSD in the same year (95% CI: 0.00021, 0.01348). No association between WLL and standardized ELA test score was found. The full regression results with coefficients and the corresponding standard errors for all six models are presented in the Additional file 1. Among the covariates, strong predictors of standardized test score include single-mom household rate, unemployment rate, student racial composition, proportion of ELL and proportion of economically advantaged students. The aim of this study is to obtain an unbiased estimate of the association between WLL and standardized test score, and collinearities among covariates are not precluded. Thus, the coefficients for covariates may be subject to collinearity-related bias.

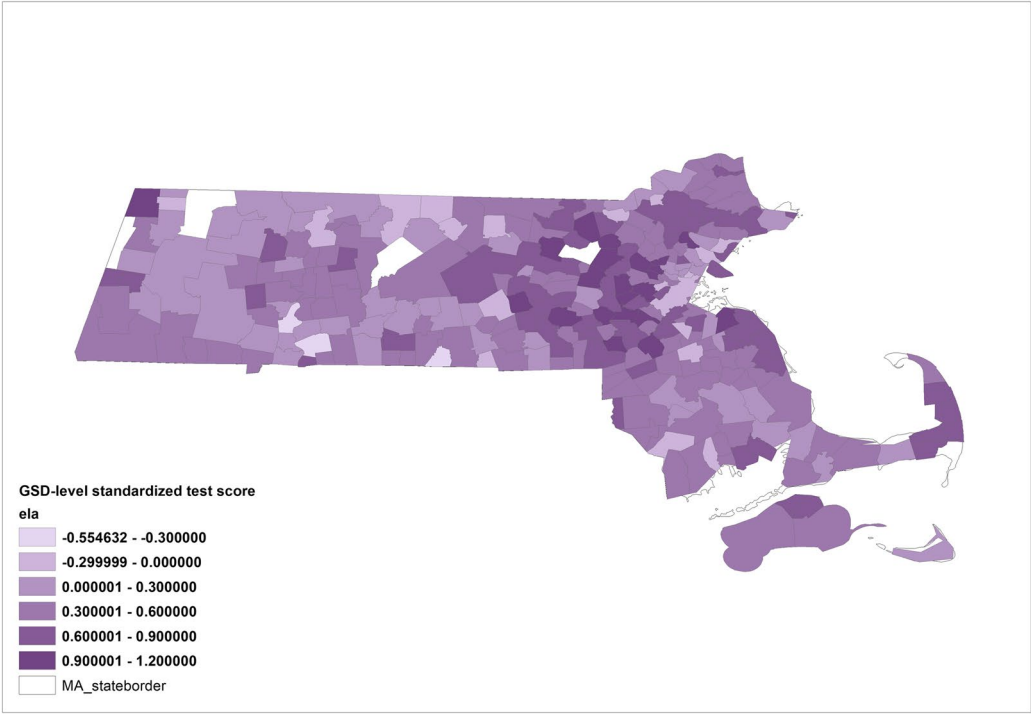
**Sensitivity analyses**

We performed several sensitivity analyses. First, we compared the final model in Table 5 Column 3 (hereafter referred to as model 3) to alternative models with cohort fixed effects but no time fixed effects, and with time fixed effects but no cohort fixed effects. Both alternative models are nested in model 3, and we used F-tests to compare these models. Model 3 is significantly better than the model with only cohort fixed effects [ $F_{(6, 6752)} = 56.969$ ;  $p < 0.001$ ], and the model with only time fixed effects [ $F_{(1227, 6752)} = 9.0629$ ;  $p < 0.001$ ]. We then used the Hausman test to compare the fixed effects model with the random effects model with the same set of covariates; the results suggested that the fixed effects model is preferred over the random effects model [ $\text{Chi-}^2_{(17)} = 586.78$ ;  $p < 0.001$ ] [55].

The standard criteria for evaluating whether adjusting for a variable or a set of variables contributes to confounding control are: (1) the variable is a predictor of the exposure, and (2) the coefficient changes after the variable is adjusted for. To evaluate whether the time fixed effects contributed to confounding control, we first calculated the Spearman’s rank correlation coefficient between WLL and year. We found that WLL in GSDs is negatively correlated with year [ $r = -0.049$ ;  $p < 0.001$ ],



(a) Math



(b) ELA

Fig. 2 GSD-level math (green) and ELA (purple) test score in Massachusetts, pooled across grades and years



**Table 5** Association of GSD-level WLL (per 5 µg/L) in CWSs with standardized test scores

	(1) $\beta$ (95% CI)	(2) $\beta$ (95% CI)	(3) $\beta$ (95% CI)
Standardized Math Test Score			
WLL (5µg/L)	-0.00967*** (-0.01631, -0.00302)	-0.00867*** (-0.01565, -0.00228)	-0.00684** (-0.01348, -0.00021)
Observations	8113	8067	8003
Standardized ELA Test Score			
WLL (5µg/L)	-0.00450 (-0.01085, 0.00185)	-0.00596* (-0.01237, 0.00045)	-0.00150 (-0.00788, 0.00488)
Observations	8112	8067	8003

All models adjusted for time fixed effects and cohort fixed effects and are weighted on total number of enrollments in each cohort-year, with 5th and 95th percentile weight truncation Model (1) is a crude model. Model (2) adjusted for urbanicity and GSD SES characteristics. Model (3) adjusted for urbanicity, GSD SES characteristics, student racial compositions and student SES characteristics. \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

thus calendar year is a predictor of WLL. We then looked at the regression coefficient for the model with only cohort fixed effects but not time fixed effects: the coefficient for a 5 µg/L change in WLL is -0.0105, which is different from the coefficient in model 3 (-0.0068). Thus, time fixed effects should be included in the model for confounding control.

We then evaluated whether cohort fixed effects contributed to confounding control. The result of a Kruskal-Wallis test [ $\text{Chi-2}_{(1250)} = 6107.7$ ,  $p < 0.001$ ] suggests that cohort is a predictor of WLLs. The regression coefficient for the model with only time fixed effects but not cohort fixed effects is 0.0122 per 5 µg/L increase in WLL, which is significantly different from the coefficient in model 3 (-0.0069). Thus, cohort fixed effects should also be included in the model for confounding control.

We also looked at models with lagged lead exposure. When the main independent variable in model 3 is replaced with WLL from the previous year, no association is found between lead exposure and standardized math test score [coefficient = -0.00047, SD = 0.00769,  $p = 0.95$ ]. This suggests that WLL is only associated with children's test scores in the same year and does not have a lag effect on the following year.

#### Effect measure modification by grades

To examine the potential effect modification by grades, we conducted a stratification analysis by grades. The stratified models applied the same covariates, fixed effects and weights as model 3 in Table 5. The regression coefficients and their respective 95% confidence intervals are shown in Fig. 3. The number of observations and the statistical power of each model was reduced after stratification, thus most regression coefficients had 95% confidence intervals covering the null. For both math and ELA, the regression coefficients appear to have a slight decreasing trend as grade increases. We then examined the statistical significance of the effect measure modification by grades using interaction terms between the WLLs and integer grades. The

effect modification by grade for math is not statistically significant [coefficient = -0.00281, standard error = 0.00205,  $p = 0.17$ ]. Because no association between WLL and ELA was found, we did not test for effect modification for ELA.

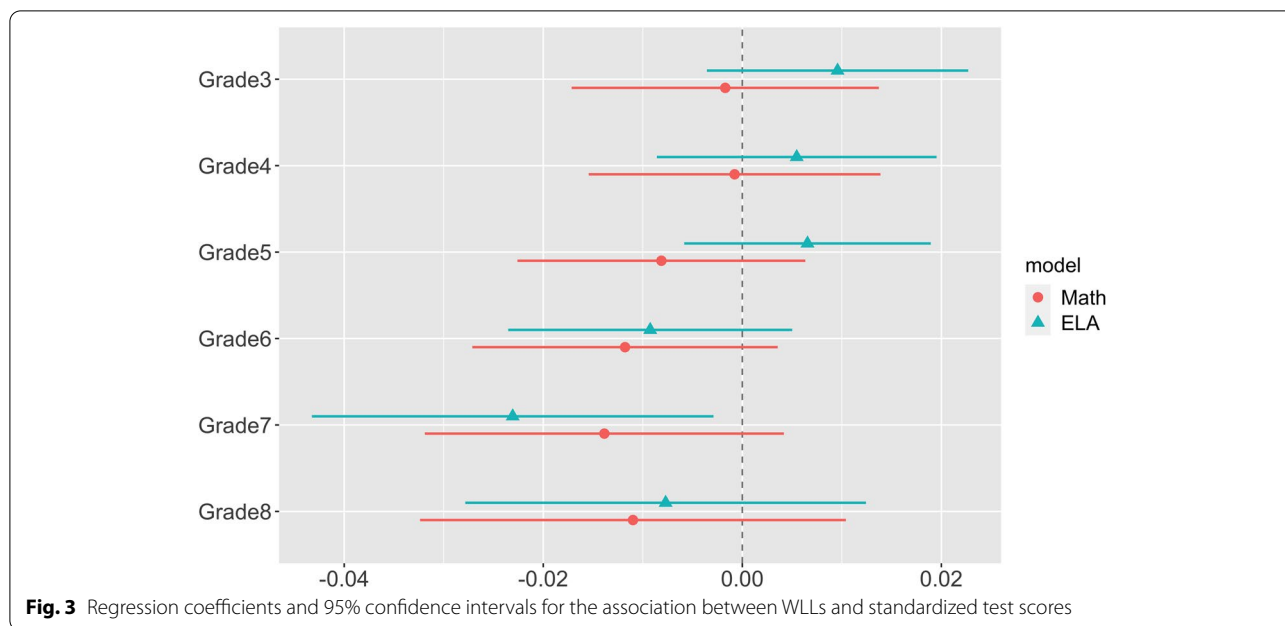
The models were adjusted for time fixed effects, cohort fixed effects, urbanicity, GSD SES characteristics, student racial composition and student SES characteristics, weighted on total number of enrollments in each cohort-year with 5th and 95th percentile weight truncation. The coefficients were presented for per 5 µg/L increase in WLLs.

#### Effect measure modification by racial compositions

We further tested the potential racial disparities of the association. A binary variable was created flagging GSD-grade-year observations with White students lower than 90%. An interaction term between this binary variable and WLL was added to model 3 and the coefficient of this term is tested. The effect measure modification by grade is statistically significant for math [coefficient = -0.0242, standard error = 0.0051,  $p < 0.01$ ]. We did not test for effect modification for ELA because no association was found between WLL and ELA test score. For cohorts with less than 90% of White students, higher WLLs are associated with a larger reduction in standardized math test scores, compared with cohort years with more than 90% of White students.

#### Discussion

Using a multivariate linear regression controlling for school district, grade level, calendar year, urbanicity, racial composition and SES, we found that a 5 µg/L higher WLL was associated with a 0.00684 SD decrease of the average math test score. Because the model has a structure similar to the DID analysis (Table 2), the exact interpretation of this association is that a 5 µg/L increase in GSD-level WLL from last year to the current year is associated with a 0.00684 SD decrease in the average Math test score of the GSD from last year to the current year. We have also found a racial disparity of this association:



the association is larger for cohorts with higher proportion of non-White students. We did not find effect modification of this association by grade.

The two-way fixed effects model (fixed effects for both unit and time) has a structure similar to the DID analysis: the DID estimator is proven to be equivalent to the two-way fixed effects model in the two group and two time period setting [51], and equivalent to a weighted two-way fixed effects model in the multi-period setting [53]. However, the causal interpretation of the association found depends on several critical assumptions. The most important assumption is the parallel trend assumption: the time-varying covariates should capture all time-varying factors other than WLL that affect the GSD-level standardized test scores. This assumption is commonly tested by looking at pre-treatment parallel trends [56]. In this study, we could not test this assumption because we cannot observe the trends of student test scores over a few years where WLLs of all GSDs remain constant, thus we cannot credit a causal interpretation of the results. Nevertheless, as analyzed in the sensitivity analysis, adding the cohort and year fixed effects controlled for unmeasured confounders that were constant within cohorts, and common trends across cohorts. For instance, lead in soil and dust is another major source of lead exposure that mainly comes from leaded paint and leaded gasoline, which were banned in the 1970s and 1990s, respectively. Thus, the levels of legacy lead deposition in outdoor soil and indoor dust from old buildings are expected to be

relatively constant or follow a common slow decreasing trend across all GSDs in MA over the study period, which are controlled for by design. The two-way fixed effects model, compared with other multivariate linear regression models, provides a closer-to-unbiased estimate of the association between WLL and children’s academic performances.

WLLs are generally low in this cohort, and well within regulatory limits. As shown in Fig. 1 and Table 4, the WLL for most GSDs are below 10 µg/L, and the median WLL level across GSDs has decreased from 4.0 µg/L to 3.3 µg/L from 2010 to 2016. However, the WLL can still be as high as 50 µg/L in some GSDs, as in years 2014, 2015 and 2016. Using estimates from our regression results in Table 5, lead exposure from WLLs of this magnitude is approximately associated with a 0.06835 SD decrease in the math test score, compared with if the lead concentration was at the lowest levels in the state (close to zero). A 50 µg/L higher WLL would move the district’s average math test score from median level in the country to the 47.3rd percentile. The association should only be interpreted on school district levels, not on an individual level.

For the same cohort of students, lead exposure and test score information were matched on years. Based on MA DEP’s lead monitoring protocols, we estimated that about 95.1% of the lead exposure data between 2010 and 2016 were from water samples collected in the Q3 monitoring period (June to October) of each year. The MCAS standardized tests for elementary and secondary

schools in MA are mostly administered during April and May [57]. Thus, for a cohort of students in this study, the lead exposure data was likely to be measured a few months after the test score outcomes. This is an important limitation in the analysis-level. We matched the exposure and outcome information in the same years assuming the ranking of the 90th percentile WLL in a GSD does not change much over a few months from April–May to the summer. This assumption is plausible because there is no required measurement of lead concentrations during these months, providing little incentive for the CWSs to change their procedures. The remaining 4.9% of the lead concentration data were measured from CWSs under standard semi-annual monitoring, where the samples were collected during the two monitoring periods Q2 and Q4 spanning across January to December, covering the months before state standard testing. The WLLs in these CWSs can be matched with the standardized test scores in the same year, with no assumption needed.

We found significant association between WLL with math test score, yet no association with ELA test score. This difference is consistent with another study that investigated the relationship between children's lead exposures and test scores in New York counties: in both bi-variate and multiple regression, the association of BLL with math test score had much higher statistical significance than BLL's association with ELA score [13]. Other studies have found BLL being associated with both math and ELA test scores with no apparent difference in statistical power [12, 14, 35, 37]. As illustrated in Fig. 3, the associations between WLL and ELA test scores were positive but not significant for children in grades 3 to 5, while negative associations have been detected among children of grades 6 to 8. The combination of slightly positive associations at younger ages and negative associations at older ages resulted in an insignificant association over all ages. The slightly positive association was found between WLL and ELA test score among younger students in this study might be due to chance given the low statistical power (caused by the relatively crude measurement and matching of the exposure), as well as the non-differential misclassification of the exposure that shifts the estimate towards the null, which will be discussed later. It does not exclude the existence of such association.

We tested for lag effects of WLL and children's test scores and did not find an association between standardized test score and WLL in the previous years. The lack of lagged association indicates that compared with the change in WLL from last year to the current year, the change in WLL from two years before to last year is not as predictive of the change in average test score from last year to the current year. This might be due to the

reduction in sample size after lagging that reduced the statistical power: the number of observations for each GSD-grade reduced from 7 to 6 after lagging. Another possibility is that there might be an acute toxicity of lead exposure. This is consistent with a study by Lanphear et al., where they tested the association between early childhood BLL, maximum BLL before IQ measurement, lifetime average BLL and concurrent BLL with child IQ deficit, and found that the concurrent BLL exhibited strongest association with IQ, indicating a potential acute effect of lead toxicity [9]. However, this does not mean that the neurotoxic effect of lead is only short-term. Studies have shown that the majority of lead in human body is stored in the bones with a very slow turnover rate [58] and can re-mobilize into the blood circulation and affect target organs in response to changes of the human body such as pregnancy [59], leading to long-term health impacts.

The implications of a detectable effect of WLLs on academic performance even at the modest levels evident in MA are significant and timely. The detected association between WLLs and math test score, although modest in magnitude, can have life-long impacts on students, such as decreased chances of college admissions and lower lifetime incomes [60, 61]. Cognitive damages associated with lead exposure impose high societal costs [62, 63] and there is good evidence of the efficacy of lead reduction efforts [64], especially for WLLs. Furthermore, EPA has recently undertaken to investigate revisions to the LCR [65]. EPA's initial proposed revisions may not reduce WLLs and hence, lead exposures [66]; however, improvements in regulatory requirements, enforcement and compliance could produce significant social benefits in the form of reduced lead damages [4, 62, 63, 67].

This study has several limitations. First, the exposure measurement was crude and ecological. WLL data was only available in its 90th percentile form with no additional information on its distribution. We rely on the assumption that PWSs with higher 90th percentile WLLs have higher lead concentrations in general. We also assume that children of grade 3 to 8 would mainly consume water from the CWS supplying their school district. To the extent that they consume water from non-community water supplies or bottled water, our exposure estimates are erroneous. This non-differential misclassification could reduce the detected effect size (closer to the null) compared to the actual effect, as well as compromise statistical power. In addition, the seasonality of the WLL sampling could not be controlled for: First, the sampling collection time was not available, thus the season of sampling of Q2 and Q4 monitoring periods could not be identified. Second, a GSD may be served by

multiple CWS with different monitoring periods, thus controlling for monitoring periods for each GSD-year observation was also not feasible. Unable to control for sampling season could potentially produce bias due to the difference in detected WLL between summer and winter, which may be influenced by temperature and disinfection practices. However, as more than 95% of the CWS-level WLLs were collected during Q3 monitoring period (summer), the magnitude of such bias is expected to be minimal. Moreover, restricting to 229 out of 294 GSDs may potentially produce selection bias, yet we found no significant difference between the included and excluded GSDs, minimizing the likelihood of selection bias due to excluded GSDs. Finally, we cannot exclude confounding by omitted time-varying covariates, although we believe there are a limited number of predictors that would be associated with WLL.

## Conclusions

Lead is a neurotoxic agent that contaminates drinking water mainly via the corrosion of plumbing components such as lead pipes and solder in the public and especially residential plumbing systems. Since EPA's 1991 LCR was issued, concentrations of lead in drinking water may have fallen, and the median WLL levels in MA, which are relatively low, have decreased from 2010 to 2016. Nonetheless, this study suggests that lead from community public water systems may still affect children's cognitive development, as reflected on their educational achievements. In school districts in MA, higher concentrations of lead in PWSs are associated with lower average math test score in the same year. Reducing math performance may have long-term impacts on children's academic success and future career options. More efforts are needed to further control lead in drinking water.

## Abbreviations

AL: Action Level; BLL: Blood Lead Level; CWS: Community Water System; DEP: Department of Environmental Protection; DID: Difference-in-difference; ELA: English Language Arts; ELL: English Language Learners; EPA: United States Environmental Protection Agency; FERPA: The Family Educational Rights and Privacy Act; GSD: Geographic School District; LCR: Lead and Copper Rule; MASSGIS: Massachusetts Bureau of Geographic Information; MCAS: Massachusetts Comprehensive Assessment System; MWRA: Massachusetts Water Resources Authority; PWS: Public Water System; SD: Standard Deviation; SEDA: Stanford Education Data Archive; SES: Socioeconomic Status; SNAP: Supplemental Nutrition Assistance Program; TAFDC: Transitional Assistance for Families with Dependent Children; WLL: Water Lead Level; WLS: Weighted Least Squares.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12889-021-12474-1>.

**Additional file 1.** Full regression results of GSD-level standardized academic test scores (math and ELA) on WLL and covariates.

**Additional file 2.** (CSV 643 kb)

## Acknowledgements

We thank Jessica Sibirski and Andrew Durham of the Massachusetts Department of Environmental Protection for providing the data of 90th percentile lead concentrations in public water systems. We thank the Stanford Center for Education Policy Analysis for maintaining the SEDA dataset and making it public.

## Authors' contributions

RL was responsible for conceptualization of the project, RL and JS were responsible for project administration and supervision, and writing-review and editing, JS was responsible for resources, WL was responsible for methodology, analysis, and writing original draft. All authors have approved the manuscript for submission.

## Funding

Funding for this study was provided by National Institute of Environmental Health Sciences (NIEHS) [grant number R30-000002]. Our funding body only provided the financial means to allow the authors to carry out the study and played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## Availability of data and materials

The datasets generated and analyzed during the current study are included in this article and its supplementary information files.

The 90th percentile lead concentration dataset of Massachusetts community water systems is maintained by the Massachusetts Department of Environmental Protection (DEP). The dataset for the most recent year (2020) is available at <https://www.mass.gov/service-details/public-water-systems-90th-percentile-lead-sampling-results>. Data from earlier years are also publicly available upon reasonable request to the Massachusetts DEP, with contact information listed in the webpage above.

The standardized math and ELA test scores dataset, GSD-level covariates dataset and GSD maps are published by Stanford University at <https://exhibits.stanford.edu/data/catalog/db586ns4974> [47].

## Declarations

### Ethics approval and consent to participate

This study did not use any individual-level human participant, material, or data. The lead in drinking water dataset is public information provided by the MA DEP with no ethical approval needed. We retrieved this dataset upon administrative permission by the Lead and Copper Rule coordinator Jessica Sibirski at the Boston office of MA DEP. The GSD-level standardized test score dataset is publicly available with no ethical approval or administrative permission needed.

### Consent for publication

Not Applicable.

### Competing interests

The authors declare that they have no competing interests.

Received: 1 September 2021 Accepted: 24 December 2021

Published online: 15 January 2022

## References

- Goldstein GW. Lead-poisoning and brain-cell function. *Environ Health Persp.* 1990;89:91–4.
- Silbergeld EK, Adler HS. Subcellular mechanisms of lead neurotoxicity. *Brain Res.* 1978;148(2):451–67.
- Sharifi AM, Baniasadi S, Jorjani M, Rahimi F, Bakshayesh M. Investigation of acute lead poisoning on apoptosis in rat hippocampus in vivo. *Neurosci Lett.* 2002;329(1):45–8.



4. Schwartz J. Low-level lead exposure and children's IQ: a meta-analysis and search for a threshold. *Environ Res.* 1994;65(1):42–55.
5. Bellinger D, Sloman J, Leviton A, Rabinowitz M, Needleman HL, Water-naux C. Low-level lead exposure and children's cognitive function in the preschool years. *Pediatrics.* 1991;87(2):219–27.
6. Canfield RL, Jusko TA, Kordas K. Environmental lead exposure and chil-dren's cognitive function. *Riv Ital Pediatr.* 2005;31(6):293–300.
7. Mazumdar M, Bellinger DC, Gregas M, Abanilla K, Bacic J, Needleman HL. Low-level environmental lead exposure in childhood and adult intel-lectual function: a follow-up study. *Environ Health.* 2011;10:24.
8. Stiles KM, Bellinger DC. Neuropsychological correlates of low-level lead exposure in school-age children: a prospective study. *Neurotoxicol Tera-tol.* 1993;15(1):27–35.
9. Lanphear BP, Hornung R, Khoury J, Yolton K, Baghurst P, Bellinger DC, et al. Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. *Environ Health Perspect.* 2005;113(7):894–9.
10. Delgado CF, Ullery MA, Jordan M, Duclos C, Rajagopalan S, Scott K. Lead exposure and developmental disabilities in preschool-aged children. *J Public Health Manag Pract.* 2018;24(2):e10–e7.
11. Chandramouli K, Steer CD, Ellis M, Emond AM. Effects of early childhood lead exposure on academic performance and behaviour of school age children. *Arch Dis Child.* 2009;94(11):844–8.
12. Zhang N, Baker HW, Tufts M, Raymond RE, Salihu H, Elliott MR. Early child-hood lead exposure and academic achievement: evidence from Detroit public schools, 2008–2010. *Am J Public Health.* 2013;103(3):e72–7.
13. Strayhorn JC, Strayhorn JM Jr. Lead exposure and the 2010 achievement test scores of children in New York counties. *Child Adolesc Psychiatry Ment Health.* 2012;6(1):4.
14. Shadbegian R, Guignet D, Klemick H, Bui L. Early childhood lead exposure and the persistence of educational consequences into adolescence. *Environ Res.* 2019;178:108643.
15. Liu J, Li L, Wang Y, Yan C, Liu X. Impact of low blood lead concentra-tions on IQ and school performance in Chinese children. *PLoS One.* 2013;8(5):e65230.
16. Blackowicz MJ, Hryhorczuk DO, Rankin KM, Lewis DA, Haider D, Lanphear BP, et al. The impact of low-level Lead toxicity on school performance among Hispanic subgroups in the Chicago public schools. *Int J Environ Res Public Health.* 2016;13(8):774. <https://doi.org/10.3390/ijerph13080774>.
17. Evens A, Hryhorczuk D, Lanphear BP, Rankin KM, Lewis DA, Forst L, et al. The impact of low-level lead toxicity on school performance among children in the Chicago public schools: a population-based retrospective cohort study. *Environ Health.* 2015;14:21.
18. Hollingsworth A, Huang M, Rudik IJ, Sanders NJ. Lead Exposure Reduces Academic Performance: Intensity, Duration, and Nutrition Matter. Working Paper Series. (NBER Working paper No.28250) [Working Paper]. In press 2020.
19. Hollingsworth A, Huang M, Rudik IJ, Sanders NJ. Lead Exposure Reduces Academic Performance: Intensity, Duration, and Nutrition Matter; 2020.
20. Reuben A, Caspi A, Belsky DW, Broadbent J, Harrington H, Sugden K, et al. Association of Childhood Blood Lead Levels with Cognitive Function and Socioeconomic Status at age 38 years and with IQ change and socioeconomic mobility between childhood and adulthood. *JAMA.* 2017;317(12):1244–51.
21. US EPA. Basic Information About Lead in Drinking Water: Environmental Protection Agency; 2021 [updated Dec 16 2021]. Available from: <https://www.epa.gov/ground-water-and-drinking-water/basic-information-about-lead-drinking-water>.
22. Brown MJ, Margolis S. In: NCEH, editor. Lead in drinking water and human blood Lead levels in the United States. Atlanta, GA: Centers for Disease Control and Prevention; 2012. p. 1–9.
23. Stanek LW, Xue J, Lay CR, Helm EC, Schock M, Lytle DA, et al. Modeled impacts of drinking water Pb reduction scenarios on Children's exposures and blood Lead levels. *Environ Sci Technol.* 2020;54(15):9474–82.
24. Hanna-Attisha M, LaChance J, Sadler RC, Champney SA. Elevated blood Lead levels in children associated with the Flint drinking water crisis: a spatial analysis of risk and public health response. *Am J Public Health.* 2016;106(2):283–90.
25. Ngueta G, Abdous B, Tardif R, St-Laurent J, Levallois P. Use of a cumulative exposure index to estimate the impact of tap water Lead concentration on blood Lead levels in 1-to 5-year-old children (Montreal, Canada). *Environ Health Persp.* 2016;124(3):388–95.
26. Levallois P, St-Laurent J, Gauvin D, Courteau M, Prevost M, Campagna C, et al. The impact of drinking water, indoor dust and paint on blood lead levels of children aged 1–5 years in Montreal (Quebec, Canada). *J Expo Sci Environ Epidemiol.* 2014;24(2):185–91.
27. Lanphear BP, Burgoon DA, Rust SW, Eberly S, Galke W. Environmental exposures to lead and urban children's blood lead levels. *Environ Res.* 1998;76(2):120–30.
28. Latham S, Jennings JL. Reducing lead exposure in school water: Evidence from remediation efforts in New York City public schools. *Environ Res.* 2022;203:111735.
29. Thomas HF, Elwood PC, Welsby E, St Leger AS. Relationship of blood lead in women and children to domestic water lead. *Nature.* 1979;282(5740):712–3.
30. US EPA. Lead and copper rule: United States Environmental Protection Agency; 2019 [updated Oct 15; cited 2020]. Available from: <https://www.epa.gov/dwreginfo/lead-and-copper-rule#rule-summary>.
31. Schock MR. Causes of temporal variability of lead in domestic plumbing systems. *Environ Monit Assess.* 1990;15(1):59–82.
32. Ngueta G, Prevost M, Deshommes E, Abdous B, Gauvin D, Levallois P. Exposure of young children to household water lead in the Montreal area (Canada): the potential influence of winter-to-summer changes in water lead levels on children's blood lead concentration. *Environ Int.* 2014;73:57–65.
33. Clay K, Portnykh M, Severnini E. The legacy lead deposition in soils and its impact on cognitive function in preschool-aged children in the United States. *Econ Hum Biol.* 2019;33:181–92.
34. Sauve-Syed, JE. Three Essays on Public Policy and Health. Dissertations - ALL. 874. 2018. Available from: <https://surface.syr.edu/etd/874>.
35. Miranda ML, Kim D, Galeano MAO, Paul CJ, Hull AP, Morgan SP. The rela-tionship between early childhood blood lead levels and performance on end-of-grade tests. *Environ Health Persp.* 2007;115(8):1242–7.
36. Reyes JW. Lead policy and academic performance: insights from Mas-sachusetts. *Harvard Educ Rev.* 2015;85(1):75–107.
37. Aizer A, Currie J, Simon P, Vivier P. Do low levels of blood Lead reduce Children's future test scores? *Am Econ J-Appl Econ.* 2018;10(1):307–41.
38. US Center for Disease Control and Prevention (CDC). Public Water Sys-tems; 2021 [updated Mar 30 2021]. Available from: <https://www.cdc.gov/healthywater/drinking/public/index.html>.
39. Massachusetts Department of Environmental Protection. 310 CMR 22: the assachusetts drinking water regulations; 2020 [updated Oct 2 2020]. Available from: <https://www.mass.gov/regulations/310-CMR-22-the-massachusetts-drinking-waterregulations>.
40. US EPA. Maximum contaminant level goals and national primary drinking water regulations for lead and copper (40 CFR Parts 141 and 142); 1998. Available from: <https://www.federalregister.gov/documents/1998/04/22/98-10713/maximumcontaminant-level-goals-and-national-primary-drink-ing-water-regulations-for-lead-and-copper>.
41. Laskey FA, Coppes DW, Estes-Smargiassi SA, Marx LM, Leone CH. Water System Master Plan. Massachusetts Water Resources Authority; 2018.
42. Massachusetts Water Resources Authority. MWRA's Water Service Com-munities; 2021 [updated Jan 21 2021]. Available from: <https://www.mwra.com/04water/html/watercommunities.html>.
43. Edwards M, Triantafyllidou S, Best D. Elevated blood lead in young children due to lead-contaminated drinking water: Washington, DC, 2001–2004. *Environ Sci Technol.* 2009;43(5):1618–23. <https://doi.org/10.1021/es802789w>. PMID: 19350944.
44. Gleason JA, Nanavaty JV, Fagliano JA. Drinking water lead and socioeco-nomic factors as predictors of blood lead levels in New Jersey's children between two time periods. *Environ Res.* 2019;169:409–16.
45. Commonwealth of Massachusetts. MassGIS (Bureau of Geographic Information); 2020. Available from: <https://www.mass.gov/orgs/massgis-bureau-of-geographic-information>.
46. Environmental Working Group. EWG's tap water database, 2019 update. 2020. Available from: <https://www.ewg.org/tapwater/>.
47. Reardon SF, Ho AD, Shear BR, Fahle EM, Kalogrides D, Jang H, et al. Stan-ford Education Data Archive (Version 3.0). 2019. Available from <http://purl.stanford.edu/db586ns4974>.



48. Code of Federal Regulations. Family Education Rights and Privacy (34 CFR Part 99). 1988. Available from: <https://www.ecfr.gov/current/title-34/part-99>.
49. Strumpf EC, Harper S, Kaufman JS. Fixed effects and difference-in-differences. In: Oakes JM, Kaufman JS, editors. *Methods in social epidemiology*, 2nd Edition. San Francisco, CA: Jossey-Bass & Pfeiffer Imprint, a Wiley brand; 2017.
50. Gangl M. Causal inference in sociological research. *Annu Rev Sociol.* 2010;36:21–47.
51. Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? *Q J Econ.* 2004;119(1):249–75.
52. Gunasekara FI, Richardson K, Carter K, Blakely T. Fixed effects analysis of repeated measures data. *Int J Epidemiol.* 2014;43(1):264–9.
53. Imai K, Kim IS. On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data. *Political Analysis.* Cambridge University Press; 2021;29(3):405–15.
54. Lu W, Hackman DA, Schwartz J. Ambient air pollution associated with lower academic achievement among US children: a nationwide panel study of school districts. *Environ Epidemiol.* 2021;5(6):e174.
55. Halaby CN. Panel models in sociological research: theory into practice. *Annu Rev Sociol.* 2004;30:507–44.
56. Abadie A. Semiparametric difference-in-differences estimators. *Rev Econ Stud.* 2005;72(1):1–19.
57. MDESE (Massachusetts Department of Elementary and Secondary Education). Statewide Testing Schedule and Administration Deadlines MCAS, MCAS Alternate Assessment, and ACCESS for ELLs: Massachusetts Department of Elementary and Secondary Education; 2022 [updated Jan 7 2022]. Available from: <https://www.doe.mass.edu/mcas/cal.html>.
58. Rabinowitz MB, Wetherill GW, Kopple JD. Kinetic analysis of lead metabolism in healthy humans. *J Clin Invest.* 1976;58(2):260–70.
59. Silbergeld EK, Sauk J, Somerman M, Todd A, McNeill F, Fowler B, et al. Lead in bone: storage site, exposure source, and target organ. *Neurotoxicology.* 1993;14(2–3):225–36.
60. Maani SA, Kalb G. Academic performance, childhood economic resources, and the choice to leave school at age 16. *Econ Educ Rev.* 2007;26(3):361–74.
61. Zax JS, Rees DI. IQ, academic performance, environment, and earnings. *Rev Econ Stat.* 2002;84(4):600–16.
62. Gould E. Childhood Lead poisoning: conservative estimates of the social and economic benefits of Lead Hazard control. *Environ Health Persp.* 2009;117(7):1162–7.
63. Landrigan PJ, Schechter CB, Lipton JM, Fahs MC, Schwartz J. Environmental pollutants and disease in American children: estimates of morbidity, mortality, and costs for lead poisoning, asthma, cancer, and developmental disabilities. *Environ Health Persp.* 2002;110(7):721–8.
64. Sorensen LC, Fox AM, Jung HJ, Martin EG. Lead exposure and academic achievement: evidence from childhood lead poisoning prevention efforts. *J Popul Econ.* 2019;32(1):179–218.
65. US EPA. National primary drinking water regulations: proposed Lead and copper rule revisions. 2019. Available from: <https://www.federalregister.gov/documents/2019/11/13/2019-22705/national-primary-drinking-water-regulations-proposed-lead-and-copper-rule-revisions>.
66. US EPA. Proposed Revisions to the Lead and Copper Rule. Published in the Federal Register March 20, 2021. Available from: <https://www.epa.gov/ground-water-and-drinking-water/proposed-revisions-lead-and-copper-rule>.
67. Grosse SD, Zhou Y. Monetary valuation of Children's cognitive outcomes in economic evaluations from a societal perspective: a review. *Children.* 2021;8(5):352.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

