The Flint, Michigan water crisis renewed concern about lead toxicity in drinking water. While lead in drinking water has been shown to negatively affect cognition among children, much less is known about its long-term consequences for late-life cognition. Using a nationally representative sample of U.S. older adults linked to historical administrative data from 1940, we find that older adults who lived as children in cities with lead pipes and acidic or alkaline water—the conditions required for lead to leach into drinking water—had worse cognitive functioning but not steeper cognitive decline. About a quarter of the association between lead and late-life cognition was accounted for by educational attainment. Within the next 10 years, American children exposed to high levels of lead during the 1970s will enter older ages. Our evidence highlights the need for stronger actions to identify interventions to mitigate long-term damage among people at high risk.

INTRODUCTION
No amount of lead is considered safe for human consumption. Lead is a neurotoxin that can cause permanent damage to the developing brains of children (1). Extensive research has documented that early-life lead exposure, even at low levels, is associated with reduced attention span and poorer academic performance among children (2–6) and lower intelligence quotient among young adults (7). Recently, this literature has expanded to link childhood lead exposure to cognitive outcomes in midlife and later adulthood; lead-exposed children were significantly associated with reduced gray matter volume in areas responsible for memory function and dementia at age 45 (8) and lower language/executive function at age 64 (9). While this line of research has been critical in identifying childhood lead exposure as an important contributing factor for brain aging, previous studies have focused on regional samples that may not be widely generalizable with short periods of follow-up. It is largely unknown whether childhood lead exposure affects trajectories of cognitive functioning throughout late adulthood in a nationally representative sample.

Individuals can be exposed to lead through multiple sources including leaded gasoline, paint, pesticides, and air pollution; however, the dominant source of lead exposure for cohorts born in the early 20th century was public drinking water (10, 11). Lead does not occur naturally in water sources. Instead, it leaches into water as it is transported through lead service lines to individual houses and buildings. The amount of lead that leaches from service lines varies by water pH, as shown in Fig. 1. Lead solubility is greatest in acidic (pH ≤6.5) or alkaline water (pH ≥8.5); these extremes can corrode or dissolve metals substantially more easily than neutral water. For cities with lead service lines, highly acidic or alkaline water leaches more lead than neutral water, while no lead can leach into drinking water in cities that did not use lead service lines regardless of water pH (12).

Service lines connecting homes to street mains in the early 20th century were commonly made of lead because of its malleability and durability (13). Although some officials were concerned about the potential toxicity of lead from water that passed through lead pipes as early as the mid-19th century (13), the federal government did not begin regulating lead until the passage of the Safe Drinking Water Act in 1974. As a result, lead concentrations in tap water in that era frequently far exceeded 0.0015 parts per billion (ppb), which is currently defined by the Environmental Protection Agency (EPA) as the maximum acceptable level in drinking water; cities in 1900 had water lead concentrations, on average, 20 to 100 times greater than the current limit (14). For example, Lowell, Massachusetts had lead concentrations of 0.1608 ppb in tap water in 1900, which exceeded 100 times the current EPA standard.

Conceptually, childhood lead exposure may shape cognitive aging through direct or indirect processes. It may directly influence cognitive functioning at older ages through “biological embedding” pathways (15), as shown in Fig. 2. The argument is that lead exposure in early life can cause early brain damage (16) that may alter gene expression (17) and elevate protein production (18), leading, in turn, to increased risk of cognitive impairment in later adulthood (19). Another hypothesized direct biological pathway is through the remobilization of lead stored in bones, which can lie dormant for decades after exposure (19, 20). Once lead is in the body, most of it stays in the bone until bone loss occurs (e.g., through osteoporosis); as a result of bone loss, some lead can reenter the blood and soft tissue. Remobilization of bone lead might, therefore, prompt a large increase in blood lead levels decades after exposure and directly exacerbate age-related cognitive decline.

Early-life lead exposure may indirectly shape cognitive aging by influencing adult achievements and experiences that subsequently matter for cognitive functioning in late life. One well-established and important indirect mechanism is through educational attainment. Lead exposure has been linked to fewer years of schooling and increased risk of high school dropout (21–23); education, in turn, is closely related to adult socioeconomic status (SES). Lower levels of education often lead to occupations that involve low mental demands and stimulation, which could reduce the brain’s capacity to sustain function amid the brain pathology and neuronal losses associated with normal aging (24). Cardiovascular health represents another pathway through which childhood lead exposure could conceivably affect late-life cognition. Lead is known to disrupt not only the central...
nervous system but also other organ functions and systems integral to brain health including the cardiovascular system (19). Among adults, occupational and residential lead exposure has been linked to hypertension, stroke, and cardiovascular disease (25, 26), which can damage blood vessels in the brain, causing lower cognitive function. Older adults with lower levels of education and poor cardiovascular health may thus enter old age with higher risk of cognitive impairment and faster decline. However, very little research has examined whether childhood lead exposure affects late-life cognition indirectly through adult education, SES, and health pathways (9, 19).

We examine the long-term association between childhood lead exposure and trajectories of cognitive change in late life among respondents to the Health and Retirement Study (HRS); we use data from the 1998–2016 biannual survey waves (27). As the longest-running aging survey of its kind, the HRS allows for more than two decades of follow-up on cognitive functioning in late life, along with financial status and cardiovascular health that we hypothesize to be important indirect pathways shaping cognition in late life. It is also one of the only publicly available nationally representative samples of U.S. older adults that has been linked to the 1940 U.S. Census. The main advantage of using these linked data is that it allows us to identify the city in which respondents lived as children, to construct early-life measures of lead exposure from municipal drinking water systems. Given that lead was a ubiquitous and poorly regulated environmental exposure for children born across much of the 20th century, it is vital to understand the potential impact that this pollutant has had on the cognitive health of these cohorts as they age.

---

**Fig. 1. Conceptual framework of life course processes through which childhood lead exposure shapes cognitive function.** Incorporating the life course perspective, we posit that childhood lead exposure may affect cognitive functioning through direct (solid line) and indirect (dashed line) pathways. The italicized font indicates demographic and childhood covariates adjusted for in all models that include age, sex, race, childhood SES, childhood health, and place of birth.

**Fig. 2. Location of cities that had lead pipes and acidic/alkaline water in 1940, the conditions required for lead to leach into municipal water.** Cities with lead pipes and acidic/alkaline water are marked on the map with red solid circles, and those with lead pipes and neutral water are marked on the map with red open circles. Cities using nonlead pipes regardless of water pH are marked on the map with gray open circles. Water pH level less than 6.5 or greater than 8.5 is considered acidic and alkaline water, respectively. Data source: The Manual of American Water-Works (30) and the U.S. Geological Survey (31).
We hypothesize that older adults who, as children, lived in cities with lead-contaminated water will have lower cognitive function and faster cognitive decline in late life compared with peers who were not exposed to lead-contaminated water. Because childhood exposures can influence adult social positions and cardiovascular health that may, in turn, affect cognitive aging, we also hypothesize that the association between childhood lead exposure and late-life cognition will be accounted for by education, adult SES, and health.

Data

We combine data from 10 waves of the HRS (1998 to 2016) and multiple sources of historical administrative data. The HRS was first conducted in 1992 on a national sample of persons born in 1931 to 1941. The following year, the HRS fielded the cohort born before 1924. In 1998, these cohorts were merged, and to create a sample representative of Americans of age 51 and older, two new cohorts born in 1924 to 1930 and 1942 to 1947 were enrolled and interviewed biannually thereafter. Among HRS respondents who were born by 1940 (n = 20,066), 9654 were successfully matched to their household records in the 1940 U.S. Census (28) that became available to the public after a mandatory 72-year waiting period (29). The matching was done on the basis of the respondents’ first and last names (including maiden name for women), age, sex, state of birth, and the names of other people living in the household in 1940 (e.g., parents and siblings). A machine learning algorithm was used to identify the correct record from among all possible records in the 1940 Census. The algorithm was trained to minimize false-positive matches and maximize the overall linkage rate. More information on the linking procedures is described in Materials and Methods.

Information on the location of cities that used lead pipes and had acidic or alkaline water in Fig. 3. Cities in the Northeast and Midwest were more likely to use lead pipes and less likely to have pH neutral water (i.e., acidic or alkaline) than those in the South and West.

Of the 9654 HRS sample members who were successfully linked to the 1940 Census, we excluded individuals who were aged over 16 to focus on those who were children or adolescents in 1940. We excluded those who were not living in cities (n = 3893) and who did not have valid piping and water pH data (n = 748). We also excluded individuals who were lost to follow-up because of death or nonresponse before the baseline survey (n = 193) and individuals with missing values on key (time-invariant) covariates (n = 215), which yielded our final analytic sample of 1089 individuals living in 398 different cities in 1940. These participants contributed 7432 observations between 1998 and 2016 (mean follow-up time = 6.8 years). We started with the 1998 Core wave because it is the first time that the sampling represents U.S. adults aged 51 and older. A chart showing the steps that we took to arrive at the analytic sample is displayed in fig. S1.

Our dependent variable is global cognitive function based on the three cognitive tests available in the HRS core survey (32): a 10-word immediate and delayed recall tests of verbal memory (0 to 20 points), a counting backward test of attention and processing speed (0 to 2 points), and a serial 7s subtraction test of working memory (0 to 5 points). A small percentage of respondents in each wave refused to participate in the cognitive tests, and to reduce sample attrition, the HRS has imputed cognitive measures for missing data (33). We used the imputed cognitive test variables released by the HRS in our analysis and standardized the summary score to have a mean of 0 and an SD of 1.

Our primary independent variable is a measure of childhood lead exposure in 1940. We identified HRS respondents who were exposed to lead as children by creating an interaction between two dichotomous variables (lead versus nonlead pipes and acidic/alkaline versus neutral water). Adult mediators included educational attainment (less than high school, high school, some college, and college or more), household income, wealth, and diagnoses of stroke, hypertension, and heart disease. We categorized both income and wealth into deciles that were treated as continuous variables in our models. All variables except education were time varying. We controlled for demographic and childhood covariates including sex, race/ethnicity, childhood SES, childhood health, and region of birth.

We used growth curve models to account for repeated observations of cognitive functioning per person (34, 35); we focus on the relationship between lead exposure and both the intercept and slope of these growth curves. Following previous work (36–38), we used ordinary least squares and logistic regression models to assess the relationship of adult mediators with childhood lead exposure. A formal mediation test using the Karlson, Holm, and Breen (KHB) method was then used to decompose the relative importance of the direct and indirect pathways from childhood lead exposure to late-life cognition (39, 40). More information on the measures and the statistical analysis is presented in Materials and Methods.

RESULTS

Table 1 presents baseline summary statistics, stratified by lead exposure status. The mean standardized global cognitive function score is 0.61, but there are substantial differences by exposure status: Respondents from cities with lead pipes and highly acidic or alkaline water had lower cognitive scores (0.35) than their peers in places with nonlead water. The chemistry of water and lead. The lead concentration varies by water chemistry. For people whose municipal water has pH values less than 6.5 or greater than 8.5, we would expect lower cognitive functioning as lead levels in their bodies would have increased as lead from water lines leaches as the result of alkaline or acidic water chemistry. However, for people whose municipal water has a pH level between 6.5 and 8.5, lead municipal water pipes are less problematic because the lead is not readily leached into the water. Adapted from (12) with permission from Elsevier.
pipes (0.59) or lead pipes and neutral water (0.68). Compared to those who were not exposed to lead in drinking water as children, the lead-exposed group had significantly higher levels of many of the risk factors that we hypothesized to be important for cognitive decline, including lower educational attainment, less household income (decile), and a higher prevalence of heart disease. The non-exposed group, while advantaged in terms of adult education and SES conditions, had a significantly higher prevalence of southern birth. The lead-exposed group was primarily born in the Northeast, had fewer female members, and far fewer Black members compared to the nonexposed group.

Results from growth curve models predicting cognitive trajectories are shown in Table 2. Results from model 1 indicate that childhood lead exposure predicts lower levels of cognitive functioning after adjustment for age, sex, race/ethnicity, childhood SES, childhood health, and region of birth. We find that individuals exposed to lead as children had, on average, 0.41 SD lower cognitive functioning at age 72 relative to those from cities with nonlead pipes and/or lead pipes (0.59) or lead pipes and neutral water (0.68).

### Table 1. Summary statistics of analytic sample at baseline, HRS (n = 1089). These statistics are shown separately by lead exposure status. We classified respondents from cities that had lead pipes and extreme levels of water pH (acidic or alkaline), the conditions required for lead to leach into drinking water, as exposed. Because, for cities that did not use lead pipes or that had neutral water, lead cannot leach into drinking water regardless of water pH, we considered the rest of respondents unexposed. We performed significant group comparisons based on chi-square test, Kruskal-Wallis test, and t test.

| Percentage (N) or mean (SD) | Exposed          | Nonexposed       | P value |
|-----------------------------|------------------|------------------|---------|
| **Full sample**             |                  |                  |         |
| Cognitive score (0–27)      | 17.29 (4.02)     | 16.10 (4.14)     |         |
| Standardized cognitive score| 0.61 (0.87)      | 0.35 (0.90)      |         |
| Lead pipes (%)              | 43.25 (471)      |                  |         |
| Acidic/alkaline water (%)   | 10.47 (114)      |                  |         |
| Age (57–93)                 | 65.49 (5.01)     | 65.52 (4.81)     |         |
| Female (%)                  | 50.69 (552)      | 45.45 (35)       |         |
| Race/ethnicity (%)          |                  |                  | P < 0.01|
| Non-Hispanic, white         | 88.89 (968)      | 97.40 (75)       | 87.38 (540) |
| Non-Hispanic, Black         | 7.99 (87)        | 2.60 (2)         | 8.74 (54) |-0.07|
| Hispanic                    | 2.30 (25)        |                  | 3.24 (20) |
| Other                       | 0.83 (9)         |                  | 0.65 (4) |
| Childhood SES (%)           |                  |                  | P = 0.63|
| 0                           | 45.55 (496)      | 40.26 (31)       | 46.76 (289) |
| 1                           | 24.24 (264)      | 31.17 (24)       | 23.95 (148) |
| 2                           | 17.26 (188)      | 18.18 (14)       | 16.83 (104) |
| 3+                          | 12.95 (141)      | 10.39 (8)        | 12.46 (77) |
| Poor childhood health (%)   | 7.07 (77)        | 10.39 (8)        | 6.31 (39) |
| Southern born (%)           | 16.53 (180)      | 2.60 (2)         | 16.18 (100) |
| Educational attainment (%)  |                  |                  | P < 0.05|
| <High school                | 12.30 (134)      | 20.78 (16)       | 12.46 (77) |
| High school                 | 39.12 (426)      | 44.16 (34)       | 37.86 (234) |
| Some college                | 20.75 (226)      | 12.99 (10)       | 22.49 (139) |
| College+                    | 27.82 (303)      | 22.08 (17)       | 27.18 (168) |
| Income (1–10)               | 6.24 (2.63)      | 5.96 (2.48)      | 6.49 (2.64) |
| Wealth (1–10)               | 6.43 (2.66)      | 6.34 (2.52)      | 6.42 (2.68) |
| Ever had stroke (%)         | 4.41 (48)        | 3.90 (3)         | 4.37 (27) |
| Ever had hypertension       | 41.05 (447)      | 33.77 (26)       | 41.42 (256) |
| Ever had heart diseases      | 15.98 (174)      | 18.18 (14)       | 15.70 (97) |

Lee et al., Sci. Adv. 8, eabn5164 (2022) 9 November 2022 4 of 10
pipes and neutral water ($\beta = -0.408, P < 0.01$). However, we did not find childhood lead exposure to be significantly associated with rates of cognitive decline.

The negative relationship between lead exposure and baseline cognition persisted with the inclusion of educational attainment in model 2. Accounting for educational attainment reduced the magnitude of the coefficient for childhood lead exposure by 34%, but the coefficient remained statistically significant ($\beta = -0.268, P < 0.05$). Higher levels of educational attainment were associated with higher levels of cognitive functioning net of childhood and demographic covariates, with respondents with high school diplomas and college or more education having, on average, 0.491 and 0.955 SD higher levels of cognitive functioning ($P < 0.001$), respectively, than those with less than high school diplomas. These results suggest that about one-third of the association between childhood lead exposure and later-life cognition is accounted for by educational attainment.

Further adjusting for income, wealth, and cardiovascular health in models 3 and 4 did not eliminate the association between lead exposure and cognitive outcomes. Individuals exposed to lead as children have 0.261 and 0.271 SD lower cognitive functioning at age 72 after adjustment for the adult SES and cardiovascular health mediators ($P < 0.05$), respectively. While each adult SES indicator was associated with higher levels of cognitive functioning, only higher wealth was associated with a slower rate of cognitive decline ($\beta = 0.001$ and 0.002 in models 3 and 4, respectively, $P < 0.05$). Model 4 reveals that the magnitude of the estimated association between childhood lead exposure and baseline cognition was on par with or exceeded the magnitude of other well-established cognitive risk factors including stroke.

| Table 2. Growth curve modeling predicting cognitive function by childhood lead exposure, adult education, and health outcomes, HRS, 1998 to 2016 ($n = 7432$ person-wave observations). All models adjust for demographic and childhood covariates including sex, childhood SES, childhood health, and region of birth. Coefficients for these demographic and childhood covariates can be found in Table S1. AIC, Akaike information criterion; BIC, Bayesian information criterion. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. |
| Fixed effects | Model 1 | Model 2 | Model 3 | Model 4 |
| Intercept | Slope | Intercept | Slope | Intercept | Slope | Intercept | Slope |
| Age (centered) | $-0.050^{***}$ | $-0.049^{***}$ | $-0.062^{***}$ | $-0.066^{***}$ |
| Age squared | $-0.001^{***}$ | $-0.001^{***}$ | $-0.001^{***}$ | $-0.001^{***}$ |
| Lead | 0.087 | 0.002 | 0.054 | 0.002 | 0.066 | 0.001 | 0.063 | 0.001 |
| Acidic/alkaline water | 0.112 | 0.009 | 0.096 | 0.012 | 0.097 | 0.011 | 0.106 | 0.012 |
| Lead X acidic/alkaline water | $-0.408^{**}$ | $-0.006$ | $-0.268^{*}$ | $-0.009$ | $-0.261^{*}$ | $-0.006$ | $-0.271^{*}$ | $-0.006$ |
| Educational attainment (<high school = reference) | | | | |
| High school | 0.491*** | 0.005 | 0.448*** | 0.006 | 0.453*** | 0.005 |
| Some college | 0.693*** | 0.003 | 0.624*** | 0.006 | 0.627*** | 0.005 |
| College+ | 0.955*** | 0.001 | 0.829*** | 0.005 | 0.830*** | 0.004 |
| Income | 0.019*** | 0.001 | 0.019*** | 0.001 | 0.019*** | 0.001 |
| Wealth | 0.024*** | 0.001* | 0.023*** | 0.002* |
| Stroke | 0.014 | 0.010** |
| Hypertension | 0.011 | 0.000 |
| Heart diseases | | | | |
| Intercept | 0.341*** | 0.389*** | 0.608*** | 0.599*** |
| Variance components | | | | |
| Variance of random intercept | 0.373*** | 0.298*** | 0.288*** | 0.283*** |
| Variance of random slope | 0.001*** | 0.001*** | 0.001*** | 0.001*** |
| Residual variance | 0.313*** | 0.313*** | 0.312*** | 0.312*** |
| Goodness of fit | | | | |
| AIC | 15258.05 | 15070.48 | 15022.62 | 15004.94 |
| BIC | 15472.37 | 15326.28 | 15306.08 | 15329.87 |
Although the point estimate for lead was larger than the point estimate for stroke ($\beta = -0.156$, $P < 0.01$), for example, the coefficient for stroke is far more precise, further research is needed to replicate this finding in other samples.

Figure 4 displays the association between childhood lead exposure and cognitive trajectories over time, based on the estimates from model 1. As we saw in Table 2, lead-exposed individuals (red solid line) initially started with lower levels of cognitive functioning, but they did not necessarily have a steeper rate of cognitive decline over time.

Overall, the results in Table 2 suggest that childhood lead exposure has little relationship with the rates of cognitive decline, that childhood lead exposure has a sizable association with baseline cognitive functioning in later life, that about a third of that association is accounted for by educational attainment, and that very little of that association is accounted for by adult SES or cardiovascular health. To more formally explore the pathways through which childhood lead exposure might affect later-life cognitive functioning, we conducted mediation analyses (tables S2 and S3). Following previous work, in models described in table S2, we first regressed the adult SES and health mediators on childhood lead exposure, controlling for age, sex, race/ethnicity, childhood SES, childhood health, and region of birth; this provides estimates of the relationship of these adult mediators with childhood lead exposure net of covariates. We found that childhood lead exposure is unrelated to most of the adult SES and health indicators except for educational attainment. For example, lead-exposed individuals had 4.663 times greater odds of receiving less than high school education, providing partial evidence for our second hypothesis. We did not find income, wealth, stroke, hypertension, and heart diseases to be associated with childhood lead exposure. Results from the KHB method in table S3 also confirm that most of the total relationship is accounted for by the direct, rather than indirect, effect of childhood lead exposure on cognitive functioning. Only 26% of the association was explained by adult pathways, and nearly all of the indirect effect of lead exposure occurred via educational attainment.

**DISCUSSION**

Beginning in 2014, residents of Flint, Michigan were exposed to high levels of lead in drinking water when a new source of public water corroded the lead service lines. The Flint water crisis has highlighted the neurodevelopmental consequences—lower cognitive abilities and poorer academic outcomes—of lead exposure in children (41, 42). However, several unresolved questions remain regarding the long-term health and cognitive consequences of childhood lead exposure. Do people exposed to lead as children have poor cognitive function at older ages? Do they experience faster cognitive decline? To what extent is the association between lead and cognition explained by adult socio-economic position and cardiovascular health? We present results from analyses of data that include historical plumbing and water pH data merged with records from a nationally representative sample of U.S. older adults to address the three questions above that have long been concerns for social and medical science researchers and policy-makers.

We find that people who lived in cities with lead-contaminated water as children had worse baseline cognitive functioning at age 72 as compared to others who did not. The conditional association between childhood lead exposure and baseline cognitive functioning from model 1 (adjusting for demographic characteristics and early-life...
confounders) was equivalent to the effect of eight additional years of age as estimated in model 1. Educational attainment appears to be a mediator in the relationship between childhood lead exposure and late-life cognitive functioning. Other mediators—income, wealth, stroke, hypertension, and heart disease—did little to mediate the relationship between childhood lead exposure and baseline cognitive functioning. However, we found no association between childhood lead exposure and rate of cognitive decline.

One possible explanation for this pattern of findings—i.e., childhood lead exposure being associated with baseline cognitive functioning but not with the pace of cognitive decline—is that the lead-exposed respondents in our sample showed a greater rate of decline before study enrollment potentially due to early brain damage. Childhood lead exposure has been found to alter brain structure (43), harm neurobehavioral development (23), and impair academic performance (3), all of which may accelerate cognitive deficits early in the life course (19). The rate of cognitive decline for the lead-exposed older adults could, therefore, be greater earlier in life but slower after age 50. Another explanation is that selective survival before study enrollment may attenuate the effects of lead on decline (44), which may bias our estimates toward the null. However, supplementary analyses of survival using Cox proportional hazards model, in which the risk of death through 2016 was estimated as a function of childhood lead exposure, showed that childhood lead exposure does not predict survival in the HRS. Regardless, in the absence of evidence of a relationship between childhood lead exposure and pace of cognitive decline, our findings have important implications for individual cognitive aging as a higher level of cognitive performance may postpone or delay the onset of cognitive impairment (45). Further investigation is needed to better understand the role of early-life environmental exposure on cognitive trajectories.

We contribute to the sparse literature by identifying childhood lead exposure as an important yet underappreciated source of early-life risk for later-life cognitive impairment (9, 19). Although exposures in childhood are increasingly being documented as predictors of life course health outcomes (46), efforts to identify early-life determinants of cognitive aging have largely focused on individual-level factors. Unexpectedly, little research has systematically examined the magnitude of environmental hazards over the life course. A major reason for this gap lies in limitations in available datasets. Investigating the effects of early-life lead exposure on cognitive aging requires data on both older adults’ cognitive performance and their childhood residential environment. However, in most aging studies, childhood environmental conditions are either measured poorly or not measured at all. A few studies have linked early-life environment to cognitive aging, but they have used aggregate-level measures (e.g., state of birth) as proxies for early-life residence (47, 48). Using multiple sources of historical contextual data at the city level, our lead exposure captures greater spatial variation in environmental effects while eliminating recall bias that influences the accuracy of estimates of early-life exposures.

There are several other advantages of studying the historical effects of lead exposure. First, lead’s toxicity in drinking water was not widely known to the public during the period of the current study. There is little evidence that city officials used information about lead’s effects on health to decide whether to use lead pipes (49), making water-borne lead exposure an exogenous factor affecting population health. Second, lead exposure via drinking water in the early and mid-20th century was uniformly distributed within cities. Studies examining the effect of lead in modern or more recent populations have documented that children from more disadvantaged backgrounds are more likely to be exposed to lead than those from advantaged background (50, 51); this is because of neighborhood-to-neighborhood variation in the ability to replace older water pipes. This makes it difficult to separate the effects of lead exposure from the effects of contextual characteristics such as neighborhood poverty and segregation. Given that the entire city population was, in most cases, exposed to lead through municipal water in earlier eras (49), historical lead exposure minimizes the possibility that estimated effects on individuals may come from selection into locations with different levels of exposure. Last, our historical measure offers an efficient tool for operationalizing the long-term consequences of early-life lead exposure; the best alternative would involve measuring levels of blood or bone lead in a large, national sample of children and then following them over 50 to 80 years to measure late-life cognition. There are some prospective cohort studies that include blood lead level measurements in childhood, but it will be several more decades before these cohorts are old enough to measure late-life cognition (7, 17).

Despite these strengths, we acknowledge several limitations of the current study. The generalizability of our finding may be limited to those cognitively intact at study enrollment. Compared to our analytic sample, those who were excluded because of the failed census linkage and missing data did have lower levels of cognition (table S9). Information on early brain structure or functioning is not available in the HRS; therefore, it is not feasible to explicitly examine cumulative biological effects through which early-life exposure to lead affects cognitive aging. Furthermore, other unobserved factors such as temperature, chlorination, flow rate, and the presence of other metal concentrations in the water are important in lead leaching (52). Considering only tap water excludes environmental exposure through other well-documented pathways, such leaded gasoline and paint. The historical measure of lead exposure may not be as precise as directly assessing lead levels in blood or bones. These limitations make service line material and water pH a crude measure, at best, of actual lead exposure, which might have biased our effect estimates toward the null. Despite this, our measure has proved useful in other population health studies to study the effect of lead on infant mortality, young adult cognition, and violent crime (10, 12, 49). Future research replicating our findings with other sources of lead exposure may be better able to estimate the effect of early-life lead exposure on health over time.

Legislative restrictions on leaded gasoline and paint are thought to have eliminated lead from communities, but residents of many U.S. cities are still exposed to high levels of lead via drinking water (53–55). Recent evidence shows that more than 40% of schools in the United States have higher than EPA-recommended levels of lead in their school tap water (56). Our findings point to the need for stronger actions in water management, corrosion control, and regulation at the state and local levels to avert future lead exposure and its enduring health consequences.

Our findings are also germane to public health concerns about American children born during the 1960s, 1970s, and 1980s who were exposed to historically unprecedented levels of lead via leaded gasoline and other sources. These cohorts had blood lead levels, on average, three times the current reference value (57), and within the next 10 to 20 years, they will enter ages at which dementia risk is heightened. McFarland et al. (58) recently estimated that 170 million Americans alive today—more than half the population—were exposed...
to high levels of lead as children. More research is clearly and urgently needed to better understand the lifelong implications of childhood lead exposure for brain aging and to identify effective interventions to mitigate lead’s long-term consequences.

**MATERIALS AND METHODS**

**1940-HRS records linkage**

The linking procedures involved the following three steps: (i) preparing and formatting data files containing HRS respondents’ identifying information, (ii) deploying machine learning algorithms to mechanically link HRS records to the 1940 U.S. Census, and (iii) hand linking records that could not be machine linked and hand-verifying a portion of those that could be (28). Because the results of machine record linkage can vary by formatting issues (e.g., whether strings are in uppercase or lowercase, whether spaces have been removed, and whether punctuation is included), data were cleaned to standardize place names and given names. For example, the strings “North Dakota” and “N. Dakota” may be treated as different strings by a computer, although they may seem the same by a human being.

The machine linkage algorithms considered the HRS respondents’ first names, last names, women’s last names in 1940, states of birth, and years of birth to locate corresponding individuals in the 1940 Census. To match the HRS respondents’ records to the 1940 Census, the population of potential matches was first identified including pairs of 1940-HRS records that displayed identical or similar characteristics on features that should be consistent over time such as year of birth. The probabilistic linking algorithm was then used to identify the correct record from among all possible records in the 1940 Census. The algorithm was trained to assign scores to pairs of 1940-HRS records to determine which single 1940 Census record is most likely to be the HRS sample member.

Last, hand linking and verification procedures were used to ensure high-quality matches and adjudicate between multiple possible matches. Other information of HRS respondents such as siblings’ and/or children’s names, parents’ years of birth, and race/ethnicity was used beyond the information that was used in the machine linking. More information on the linking procedures has been described in detail elsewhere (28).

**Measures**

**Cognitive functioning**

For the immediate and delayed word recall tests, respondents were given a list of 10 nouns by interviewers and asked to recall as many words as possible from the list in any order. After approximately 5 min of asking other survey questions, the respondent was asked to recall the nouns previously presented as part of the immediate recall task. We counted the total number of immediate and delayed words that were recalled correctly (0 to 20 points). For the counting backward test, respondents were asked to count backward for 10 continuous numbers beginning with the number 20 as quickly as possible. We scored 2 points if answered correctly on the first try; 1 point if correctly answered on the second try; 0 if incorrect on the first or second try (0 to 2 points). For a serial 7s subtraction test, respondents were asked to subtract 7 from 100 and continue subtracting 7 from each subsequent number for a total of five trials. We counted the number of correct subtractions among the five trials (0 to 5 points). The summary score ranged from 0 (severely impaired) to 27 (high functioning) (32). Although additional cognitive status variables are assessed in the HRS, they are only administered to respondents aged 65 and older. Hence, we included these three measures administered to all sample members.

**Childhood lead exposure**

We use historical data about municipal service lines and water chemistry as a proxy measure of lead exposure via lead-contaminated drinking water. For cities that did not use lead service lines, no lead can leach into drinking water regardless of water pH. However, for cities with lead service lines, highly acidic (pH ≤6.5) or alkaline (pH ≥8.5) water leaches more lead than neutral water (12).

**Adult SES and cardiovascular outcomes**

Adult SES was measured using educational attainment (less than high school, high school, some college, and college or more) and household income and wealth. Both income and wealth were drawn from the RAND Income and Wealth Imputation data (version V1). Some respondents have zero income and negative wealth. We categorized income and wealth into deciles that were treated as a continuous variable in our models. Stroke, hypertension, and heart disease (diagnosed = 1 and nondiagnosed = 0) were considered as cardiovascular outcomes. We coded 1 if respondents reported that their doctors told them that they had a stroke, hypertension (or high blood pressure), or a heart disease (a heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems). All variables except education were time varying.

**Covariates**

Demographic covariates include sex (female = 1) and race/ethnicity (non-Hispanic white, non-Hispanic Black, Hispanic, and other). We also controlled for childhood covariates including childhood SES, childhood health, and region of birth. We measured an index of childhood SES that included five items from the core survey: father’s education (less than 8 years = 1), mother’s education (less than 8 years = 1), self-reported financial situation (financially poor = 1), moved because of financial difficulty (yes = 1), and received help from relatives because of financial difficulty (yes = 1). The Cronbach’s alpha was 0.71, indicating good reliability among the items. Following previous work (37, 38), we imputed missing data for mother’s and father’s education as less than 8 years as respondents in the HRS missing data on parental education have economic and health variables similar to those whose parents had less than 8 years of education. Few respondents experienced four or more adversities, so we collapsed the index to range from zero to three or more. Childhood health was coded 1 if respondents reported in the core that they had “poor” or “fair” health. We compared individuals who were born in southern states (=1) to those who were born in non-southern states (=0).

**Analytic strategy**

We first estimated summary statistics of key variables used in our study at baseline. These statistics are shown separately by lead exposure status. We classified respondents from cities that had lead pipes and extreme levels of water pH (acidic or alkaline), the conditions required for lead to leach into drinking water, as exposed. Because, for cities that did not use lead pipes or that had neutral water, lead cannot leach into drinking water regardless of water pH, we considered the rest of respondents unexposed. We performed significant group comparisons on the basis of chi-square test, analysis of variance (ANOVA), or Kruskal-Wallis test.

Next, we fit a series of nested models to assess the association between childhood lead exposure and cognitive trajectories. We first fit model 1 predicting the long-term association between childhood lead exposure and trajectories of cognitive function. We
then accounted for educational attainment in model 2 because education is prior to income and wealth attainment (38). Model 3 considered income and wealth. Model 4 further adjusted for adult cardiovascular health conditions. Accounting for adult SES and health indicators in this sequential manner allowed us to investigate to what extent the association between childhood lead exposure and cognitive functioning observed in model 1 is mediated by the adult SES and health pathways (38). All models were controlled for covariates including age, sex, race/ethnicity, childhood SES, childhood health, and region of birth. The unconditional model (model 0) estimating the association between childhood lead exposure and late-life cognitive trajectories without childhood covariates can be found in table S1.

Following procedures previously described in (38), we fit these models using growth curve models. This analytical framework accounts for partially missing (or unbalanced) data using maximum likelihood and performs equally well or better than multiple imputation methods (34, 35). The mean number of observations per respondent was 6.8. Time was modeled by age in the current study and was centered on its grand mean at age 72. Preliminary analyses showed a nonlinear relationship between age and the outcome, so we added age squared to all models. All independent variables and covariates were interacted with age to test for differences in the rate of change in cognitive functioning. The model is expressed formally as

$$
\text{Cognitive functioning}_{ij} = \gamma_{10} + \gamma_{11}\text{Age}_{ij} + \gamma_{20}\text{Age}_{ij}^2 + X_{ij}\beta + (X_{ij} \times \text{Age}_{ij})\lambda + \zeta_{0i} + \zeta_{1i}\text{Age}_{ij} + \epsilon_{ij}
$$

where subscripts i and j index the individual and person-wave observation, respectively. i represents individual, and j represents measurement occasion, which, in this case, is survey wave. \(\gamma_{10}\) denotes the fixed intercept, \(\gamma_{11}\) and \(\gamma_{20}\) indicate fixed effects for the linear and quadratic terms of age, while \(\beta\) is the vector of coefficients associated with the vector of covariates \(X\) (e.g., \(X\) can be either time invariant or time varying; the coefficients for time-invariant childhood variables are the same at all \(j = 1, 2, 3, \ldots, 10\)). \(\lambda\) is the vector of coefficients associated with the vector of covariates and their interaction with age. The \(\zeta_{0i}\) and \(\zeta_{1i}\) terms represent normally distributed random effects for the intercept and linear term of age, respectively (these were allowed to covary in the model). All models were estimated using mixed in Stata 17.

Mediation analyses were carried out in two steps. First, following previous work (36–38), we ran a series of models with adult SES and health outcomes as a function of childhood lead exposure, net of demographic, and childhood covariates. We chose to analyze adult-SES and health indicators with MRI measurements of structural brain integrity in midlife. All models were estimated using maximum likelihood and performs equally well or better than multiple imputation methods (34, 35). The mean number of observations per respondent was 6.8. Time was modeled by age in the current study and was centered on its grand mean at age 72. Preliminary analyses showed a nonlinear relationship between age and the outcome, so we added age squared to all models. All independent variables and covariates were interacted with age to test for differences in the rate of change in cognitive functioning. The model is expressed formally as

$$
\text{Cognitive functioning}_{ij} = \gamma_{10} + \gamma_{11}\text{Age}_{ij} + \gamma_{20}\text{Age}_{ij}^2 + X_{ij}\beta + (X_{ij} \times \text{Age}_{ij})\lambda + \zeta_{0i} + \zeta_{1i}\text{Age}_{ij} + \epsilon_{ij}
$$

where subscripts i and j index the individual and person-wave observation, respectively. i represents individual, and j represents measurement occasion, which, in this case, is survey wave. \(\gamma_{10}\) denotes the fixed intercept, \(\gamma_{11}\) and \(\gamma_{20}\) indicate fixed effects for the linear and quadratic terms of age, while \(\beta\) is the vector of coefficients associated with the vector of covariates \(X\) (e.g., \(X\) can be either time invariant or time varying; the coefficients for time-invariant childhood variables are the same at all \(j = 1, 2, 3, \ldots, 10\)). \(\lambda\) is the vector of coefficients associated with the vector of covariates and their interaction with age. The \(\zeta_{0i}\) and \(\zeta_{1i}\) terms represent normally distributed random effects for the intercept and linear term of age, respectively (these were allowed to covary in the model). All models were estimated using mixed in Stata 17.

Mediation analyses were carried out in two steps. First, following previous work (36–38), we ran a series of models with adult SES and health outcomes as a function of childhood lead exposure, net of demographic, and childhood covariates. We chose to analyze adult-SES and health indicators with MRI measurements of structural brain integrity in midlife. All models were estimated using mixed in Stata 17.

Mediation analyses were carried out in two steps. First, following previous work (36–38), we ran a series of models with adult SES and health outcomes as a function of childhood lead exposure, net of demographic, and childhood covariates. We chose to analyze adult-SES and health indicators with MRI measurements of structural brain integrity in midlife. All models were estimated using mixed in Stata 17.

Mediation analyses were carried out in two steps. First, following previous work (36–38), we ran a series of models with adult SES and health outcomes as a function of childhood lead exposure, net of demographic, and childhood covariates. We chose to analyze adult-SES and health indicators with MRI measurements of structural brain integrity in midlife. All models were estimated using mixed in Stata 17.

Mediation analyses were carried out in two steps. First, following previous work (36–38), we ran a series of models with adult SES and health outcomes as a function of childhood lead exposure, net of demographic, and childhood covariates. We chose to analyze adult-SES and health indicators with MRI measurements of structural brain integrity in midlife. All models were estimated using mixed in Stata 17.

Mediation analyses were carried out in two steps. First, following previous work (36–38), we ran a series of models with adult SES and health outcomes as a function of childhood lead exposure, net of demographic, and childhood covariates. We chose to analyze adult-SES and health indicators with MRI measurements of structural brain integrity in midlife. All models were estimated using mixed in Stata 17.

Mediation analyses were carried out in two steps. First, following previous work (36–38), we ran a series of models with adult SES and health outcomes as a function of childhood lead exposure, net of demographic, and childhood covariates. We chose to analyze adult-SES and health indicators with MRI measurements of structural brain integrity in midlife. All models were estimated using mixed in Stata 17.
22. J. P. Nilsson, The long term effects of Early Childhood Lead Exposure: Evidence from the Phase-out of Lead Gasoline (Institute for Labour Market Policy Evaluation/IFAU Work. Pap., 2009), pp. 1–59.

23. H. L. Needleman, A. Schell, D. Bellinger, A. Leviton, E. N. Allred, The long-term effects of exposure to low doses of lead in childhood. *N. Engl. J. Med.* 322, 83–88 (1990).

24. Y. Stern, Cognitive reserve. *Neuropsychologia* 47, 2015–2028 (2009).

25. A. Menke, P. Munterm, V. Batuman, E. K. Silbergeld, E. Guallar, Blood lead below 0.48 microg/L (10 microg/dl) and mortality among US adults. *Circulation* 114, 1388–1394 (2006).

26. A. Navas-Acien, E. Guallar, E. K. Silbergeld, S. J. Rothenberg, Lead exposure and cardiovascular disease—A systematic review. *Environ. Health Perspect.* 115, 472–482 (2007).

27. A. Sonneg, J. D. Faul, M. B. Ostfeld, K. M. Langa, J. W. Phillips, D. R. Weir, Cohort profile: The health and retirement study (HRS). *Int. J. Epidemiol.* 43, 576–585 (2014).

28. J. R. Warren, F. Pfeffer, J. Helgertz, D. Xu, Linking 1940 U.S. Census Data to the Health and Retirement Survey: Technical Documentation, in *HRS Documentation Report* (Release 1. Survey Research Center, Institute for Social Research, University of Michigan, 2020).

29. S. Ruggles, C. A. Fitz, E. Roberts, Historical census record linkage. *Annu. Rev. Social. Sci.* 44, 19–37 (2018).

30. M. N. Baker, *The Manual of American Water-Works* (Engineering News, 1897), vol. 4.

31. E. W. Lohr, S. K. Love, *The Utility of Public Water Supplies in the United States, 1952* (US Government Printing Office, 1954).

32. E. M. Crimmins, J. K. Kim, K. M. Langa, D. R. Weir, Assessment of cognition using surveys and neuropsychological assessment: The health and retirement study and the aging, demographics, and memory study. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 66, 1162–1171 (2011).

33. G. G. Fisher, H. Hassan, J. D. Faul, W. L. Rodgers, D. R. Weir, Health and Retirement Study imputation of cognitive functioning measures: 1992–2014 (final release version) data description, in *Ann Arbor, MI: University of Michigan, Survey Research Center (Survey Research Center, University of Michigan, 2017)*; https://hrs.isr.umich.edu/sites/default/files/biblio/COGIMPdd.pdf.

34. P. J. Curran, K. Obedat, D. Losardo, Twelve frequently asked questions about growth curve modeling. *J. Cogn. Dev.* 11, 121–136 (2010).

35. J. D. Singer, J. B. Willett, *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence* (Oxford Univ. Press, 2003).

36. K. F. Ferraro, M. H. Schafer, L. R. Wilkinson, Childhood disadvantage and health problems in middle and later life: Early imprints on physical health? *Am. Sociol. Rev.* 81, 107–133 (2016).

37. Y. Luo, L. J. Waite, The impact of childhood and adult SES on physical, mental, and cognitive well-being in later life. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 60, 593–5101 (2005).

38. H. Lee, L. H. Ryan, M. B. Ostfeld, J. Smith, Multigenerational households during childhood and trajectories of cognitive functioning among U.S. older adults. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 76, 1161–1172 (2021).

39. K. B. Karlson, A. Holm, R. Breen, Comparing regression coefficients between same-sample nested models using logit and probit. *Sociol. Methodol.* 42, 286–313 (2012).

40. R. Breen, K. B. Karlson, A. Holm, Total, direct, and indirect effects in logit and probit models. *Sociol. Methods Res.* 42, 164–191 (2013).

41. S. Trejo, G. Yeomans-Maldonado, B. Jacob, *The Psychosocial Effects of the Flint Water Crisis on School-Age Children* (National Bureau of Economic Research, 2021), p. w29341.

42. S. Zheng, S. L. Bishop, T. Ceja, M. Hanna-Attisha, K. LeWinn, Neurodevelopmental profiles of preschool-age children in Flint, Michigan: A latent profile analysis. *J. Neurodev. Disord.* 13, 29 (2021).

43. K. M. Cecil, J. C. Brubaker, C. M. Adler, A. Reuben, Half of US population exposed to adverse lead levels in early childhood. *Proc. Natl. Acad. Sci.* 119, e2118631119 (2022).

44. U. Kohler, K. B. Karlson, A. Holm, Comparing coefficients of nested nonlinear probability models. *SJ* 11, 420–438 (2011).

45. J. L. Saenz, B. Downton, M. A. Garcia, R. Wong, Cognition and context: Rural-urban differences in cognitive aging among older mexican adults. *J. Aging Health* 30, 985–986 (2018).

46. H. Lee, M. Schafer, Are positive childhood experiences linked to better cognitive functioning in later life? Examining the role of life course pathways. *J. Aging Health* 33, 217–225 (2021).

Acknowledgments: We thank E. Crimmins and J. Alshire and the University of Southern California Center for Biodemography and Population Health for helpful feedback on an earlier draft. We also thank the late W. Troesken whose work has inspired us to continue to work on life after lead. An earlier version of this paper was presented as an oral paper at the 2021 the Interdisciplinary Association for Population Health Science. It was also presented and received a Poster Award at the 2022 Population Association of America. Funding: The effort to link the HRS to the 1940 Census was supported by the National Institute on Aging (NIA) via grant R01AG050300. Support has also come from the Minnesota Population Center, which receives core funding (P2CHD041023) from the Eunice Kennedy Shriver National Institute for Child Health and Human Development (NICHD). H.L. was supported by an NIA K99 Pathway to Independence Award (K99AG071834). M.W.L. was supported by a training grant from the NICHD (T32HD095134). Errors and omissions are the responsibility of the authors. Author contributions: Conceptualization: H.L., M.W.L., J.R.W., and J.F. Methodology: H.L., M.W.L., J.R.W., and J.F. Investigation: H.L. and M.W.L. Visualization: H.L. Funding acquisition: H.L. and M.W.L. Supervision: J.R.W. and J.F. Writing—original draft: H.L. and M.W.L. Writing—review and editing: H.L., M.W.L., J.R.W., and J.F. Competing interests: The authors declare that they have no competing interests. Data and materials availability: The linked data between the HRS and the 1940 Census used in this paper are available on data enclave through the Michigan Center on the Demography of Aging (MCDA) (https://hrs.isr.umich.edu/data-products/ restricted-data/available-products). All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 1 December 2021
Accepted 20 September 2022
Published 9 November 2022
10.1126/sciadv.abn5164