Predicting Lead Water Service Line Locations for Improving Mitigation Strategies

Mike Blackhurst

Co-Director, Urban and Regional Analysis

5/9/2018

University of Pittsburgh

University Center for Social & Urban Research (UCSUR)

Summary

Lead water laterals are the primary source of lead in drinking water in the U.S. Municipalities and public health agencies are typically given broad discretion on how to protect their constituents from lead in drinking water. However, effective mitigation and prevention strategies are often complicated by missing information describing the locations of lead laterals, which, in turn, makes it difficult to understand who is at risk of lead exposure. As a result, the limited resources available for mitigation and prevention may not be allocated in the most cost effective manner.

In this study, data describing the Pittsburgh Water and Sewer Authority's (PWSA) drinking water distribution system were merged with property data to develop predictive models of lead service line locations. Following expected historical development patterns, the lateral material at neighboring properties was the best predictor of lead laterals. The date of lateral installation and date of inspection were also good predictors. The year a property was built provided acceptable prediction. The model was 78% accurate in predicting lead laterals identified through recent inspections. Using a combination of recent inspections and predictions, this study estimates that approximately 39,000 lead laterals are in PWSA's service area.

Lateral replacement costs specific to each property in PWSA were estimated using geospatial data describing building locations, parcel boundaries, and street curb locations. The average length and estimated replacement cost of lead laterals in PWSA were estimated to be 26 feet (q1 = 12 feet; q3 = 31 feet) and $7,200 (q1 = $5,200; q3 = $8,900), respectively. The total cost of replacing lead laterals is estimated to be $290 million with a range of $190 million to $390 million.

Replacement costs were merged with demographic data at the Census block group and neighborhood levels to develop criteria for prioritizing prevention (lead laterals per child under age 5) and lateral replacements ($ per child under age 5 protected). Excluding neighborhoods without any lead laterals (n = 3), neighborhood variation in the incidence of lead laterals per child is estimated to vary from 0.025 to 277 (mean = 16). Neighborhood variation in replacement cost effectiveness is estimated to vary from $17,000 to over $100,000 per child protected (under age 5), with a mean cost effectiveness of $350,000 per child. These results can help prioritize scarce resources for prevention and mitigation by identifying areas where children are mostly likely exposed and the costs of mitigation are relatively low.

Results are published in map format below, and the respective model results are available for download here (https://mblackhurst.carto.com/builder/458f3aa9-4b5b-4a23-8d6a-4d3a1e8651f9/embed). Results are subject to change pending updated data and input from members of the Pittsburgh community.

1. Data sources

Data Describing Property and Water Laterals

Property characteristics were provided by Allegheny County (2017a).Property characteristics include lot size, building characteristics, recent sales information, designated use (e.g., single family), and an evaluation of the property condition performed as part of assessing the property for tax purposes.

PWSA (2018) provided addresses of customer accounts, the customer class (residential or commercial), the services provided (water and/or sewer), and varied historical data describing service laterals. Where data are available, laterals are described by their diameter, inspection results (date and materials), installed material and date, and a notes field. The lateral data were consolidated into a two variables, one indicating any historical presence of lead (historically present = 1; historically not present = 0) and one describing any indication that a lead lateral had been replaced (replaced = 1; not replaced = 0). PWSA has not perfectly maintained their historical data, so many records are missing and some inaccuracies are expected.

Assuming residential properties are served by a lateral with a maximum diameter of 1 inch, the PWSA data reflect approximately 70,000 active residential accounts associated with 95,000 properties. Where a PWSA account is associated with more than one address, the account is most often designated as a multi-family unit. Models predicting lead laterals were prepared at the address level, allowing for repeated observations of unique laterals. Lateral inventories are by PWSA account.

The PWSA data also include 17,500 inactive accounts. The inactive accounts were retained for the purpose of preparing models predicting the historical lateral material, then triaged in subsequent estimates of lead lateral inventories.

PWSA accounts where joined to the County property data by address or geolocation. Approximately 650 PWSA accounts (22 of which historically had lead) where geolocated outside of PWSA's service area, and these accounts were dropped from all analyses. The neighborhood and block group where imputed using the street address for about 50 accounts; 550 accounts

could not be geolocated.

Table 1 summarizes the sample used for predicting the lateral material. As indicated in Table 1, a total of 1,573 recently inspected laterals were set aside as a validation sample.

Table 1: Summary of sample used to predict the probability that any given residential property in Pittsburgh, PA was historically served by a lead water lateral.

	used to train model		**used to validate model**
statistic	**active accounts**	**inactive accounts**	**active accounts**
total count	***68,816***	***17,512***	***1,573***
records reporting install year	***42,061***	***8,459***	***925***
mean install year	1944	1938	1932
earliest install year	1910	1910	1910
latest install year	2017	2017	2017
records reporting inspect year	***23,147***	***13,573***	***508***
mean inspect year	1938	1936	1933
earliest inspect year	1902	1904	1906
latest inspect year	1997	2016	1992
records reporting year built	***62,903***	***4,362***	***1,521***
mean year built	1919	1921	1915
earliest year built	1800	1830	1848
latest year built	2016	2016	2016
records reporting historical lateral material	***49,628***	***14,824***	***1,573***
records indicating lead historically present	24,833	10,116	1,214
records reporting replacements	***1,677***	***2***	***1,368***
records indicating lead lateral replaced	191	0	129

Figure 1 shows the historical presence of lead laterals at up to twenty neighboring properties by street address. If a property is indicated as historically having lead, more neighbors are also likely to have lead laterals. In contrast, if a property has no historical evidence of lead, fewer neighbors have had lead laterals. Cumulative counts of neighboring lead laterals were used in predicting the historical presence of lead.

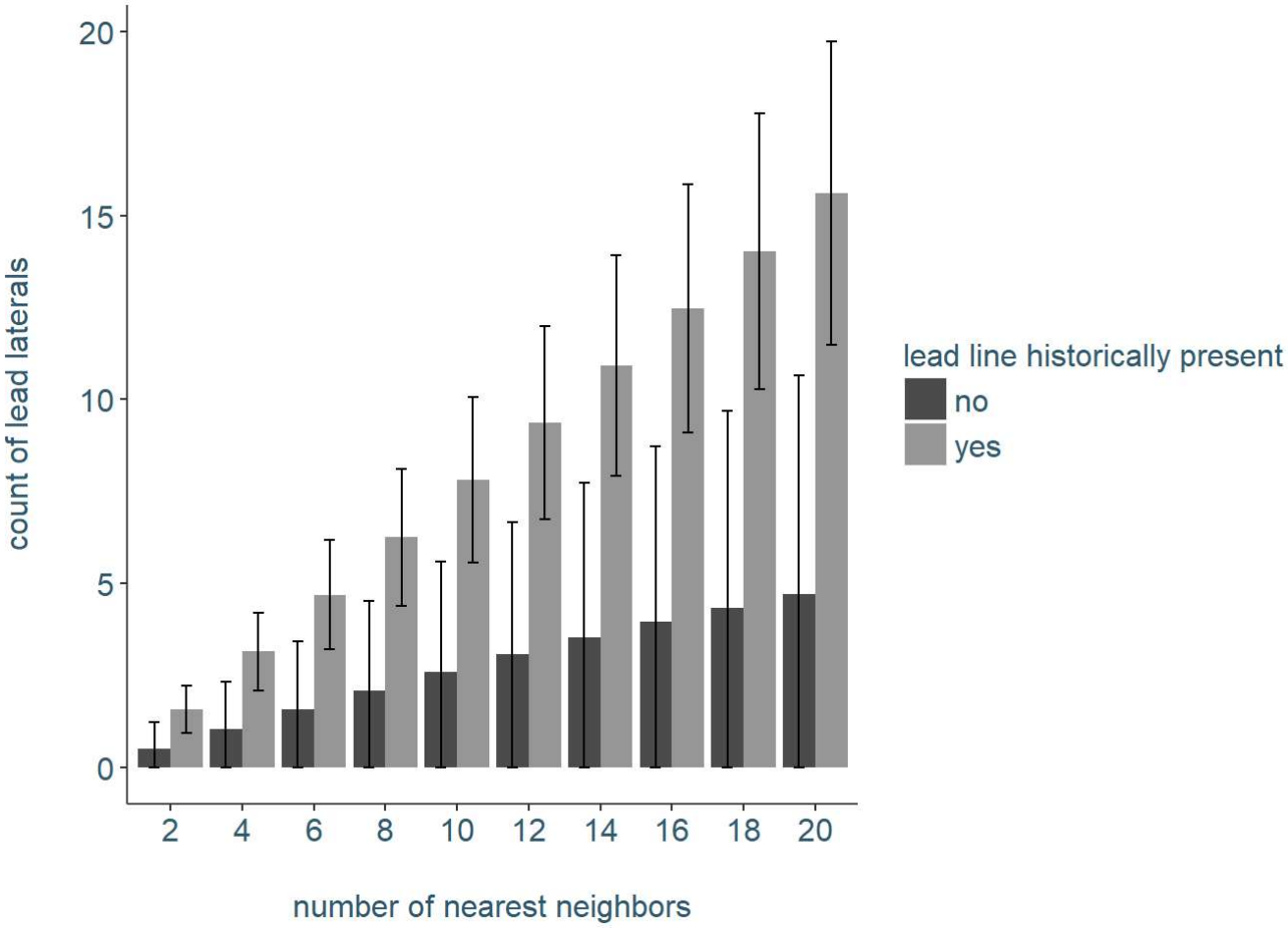

Figure 1: Counts of lead laterals at up to twenty closest neighbors by street address for Pittsburgh, PA homes indicated as historically having or not having a lead water lateral.

Historically present laterals may have been replaced. The few records with data indicating a lead lateral has been replaced suggest roughly 10.5% (320/3,047) of lead laterals have been replaced, which is indicated by a historical marker for lead and a more recent inspection indicating non-lead material. No correlations between replacements and other property characteristics were found, suggesting replacements are largely random. Thus, a 10.5% probability of replacement was uniformly applied when estimating counts of existing lead laterals.

Data Describing Lateral Replacement Costs

Street curb and building footprints were provided by the City of Pittsburgh (2015) and Allegheny County (2018), respectively. For each property, the curb setback was estimated by finding the minimum straight line distance between the largest building (removing smaller auxiliary buildings) and the curb. The lateral replacement length was assumed to be the straight line distance or a minimum of four feet. Where multiple properties are associated with a PWSA account, the median, low, and high setbacks were assumed equally probable.

Base case, low, and high unit replacement costs of $300, $200, and $500 per foot of lateral were assumed equally probable. These estimates are much lower than those currently used by PWSA, which is around $1,000 per foot assuming an average lateral length of 20 feet [Hopey 2018]. However, the cost estimates used here are more consistent with casually understood replacement cost estimates.

Geographic and Demographic Data

Neighborhood and block group spatial boundaries were provided by the City of Pittsburgh (2016) and the U.S. Census Bureau (2017), respectively. Where PWSA's service area split block group boundaries, census data were proportioned into the PWSA service area by counts of residential parcels [Allegheny County 2017a&b]. Block group data were aggregated by neighborhood. Where neighborhoods split block groups, census data were proportioned by counts of PWSA accounts.

2. Models of Lead Laterals

Logit models predicting the presence of lead given historical lateral inspection and installation dates, property and lot characteristics, lateral material at neighboring properties, and spatial location were explored. Models were diagnosed for linearity, predictive capability (McFadden's psuedo-R2), fit (Akaike Information Criterion), and sample coverage. Table 2 summarizes the logit model results.

As indicated in Table 2, not all covariates were available at each property for predictions. At each residential property in PWSA's service area, the probability of lead being historically present was predicted using the best model (highest psuedo-R2) where covariates are recorded for the property. As an example, consider a property reporting values for the lateral inspection year ("year.inspect") and counts of lead laterals at the nearest four neighbors ("neighb4"). The model "Prob(historical.lead) ~ neighb4" was used for predictions given its higher psuedo-R2. Applying the predictions for the best available model at each property to the validation sample indicates an accuracy and AUC (area under the receiver operating characteristic curve) of 78% and 80%, respectively.

To estimate whether lead is currently installed, a constant replacement probability of 10.5% was applied to the predicted probability that lead was historically present. As described below, lead laterals are assumed currently present where this joint probability is greater than or equal to 0.5.

Lead present if Prob(historical.lead)(1 - Prob(replaced)) >= 0.5*

Lead not present if Prob(historical.lead)(1 - Prob(replaced)) < 0.5*

Table 2: Summary of logit models predicting the probability any given property has a lead line. The sample count columns are not balanced because the availability of covariate values is not mutually exclusive. The covariates "neigh_x" indicate counts of lead laterals within x neighboring properties.

Model	Count address	McFadden psuedo-R2	Marginal addresses modeled	Cumulative addresses modeled
Prob(historical.lead) ~ year.install + neighb_4 + neighborhood + condition	34,841	0.793	NA	34,841
Prob(historical.lead) ~ neighb_14	32,038	0.791	11,760	46,601
Prob(historical.lead) ~ neighb_8	43,364	0.701	11,326	57,927
Prob(historical.lead) ~ neighb_4	55,981	0.587	12,617	70,544
Prob(historical.lead) ~ year.inspect	46,117	0.528	10,827	81,371
Prob(historical.lead) ~ neighb_2	65,987	0.462	26,053	107,424
Prob(historical.lead) ~ period	53,501	0.426	0	121,431
Prob(historical.lead) ~ year.built	53,501	0.394	0	121,431
Prob(historical.lead) ~ block.group	78,956	0.162	0	121,431
Prob(historical.lead) ~ neighborhood	78,743	0.130	0	121,431

3. Uncertainty Modeling

For each property predicted to have a lead line, randomly drawn samples of 1,000 potential lateral lengths, unit costs, and replacement costs were prepared, and the resulting samples were summarized by Census block group.

For demographic data (counts of child, households, and income), the Census reports a single estimate and a measurement of error (MOE), which is defined as 1.645 times the standard error. The distribution of the Census data was assumed to be zero-truncated normal, and 1000 random draws were taken for each demographic variable. The demographic data were then merged with the lead inventory and replacement costs estimated by block group and neighborhood.

4. Results

The total estimated lead lateral inventory is approximately 39,000, which is 56% of all residential PWSA customers. The incidence of lead laterals for child under the age of 5 varies by neighborhood from 0.025 to 277, with a mean incidence of 16 laterals per child. Figure 2 shows neighborhood variation in both children per household and the incidence of lead laterals per PWSA customer. The neighborhood of Beltzhoover has the highest incidence of lead at nearly 60%, but has less than the average number of children per household. Neighborhoods in the upper right quadrant of Figure 2 have a higher than average incidence of lead laterals and children per household. All else equal, prevention strategies, such as education and outreach efforts, in these neighborhoods are more likely to lead to awareness and mitigation.

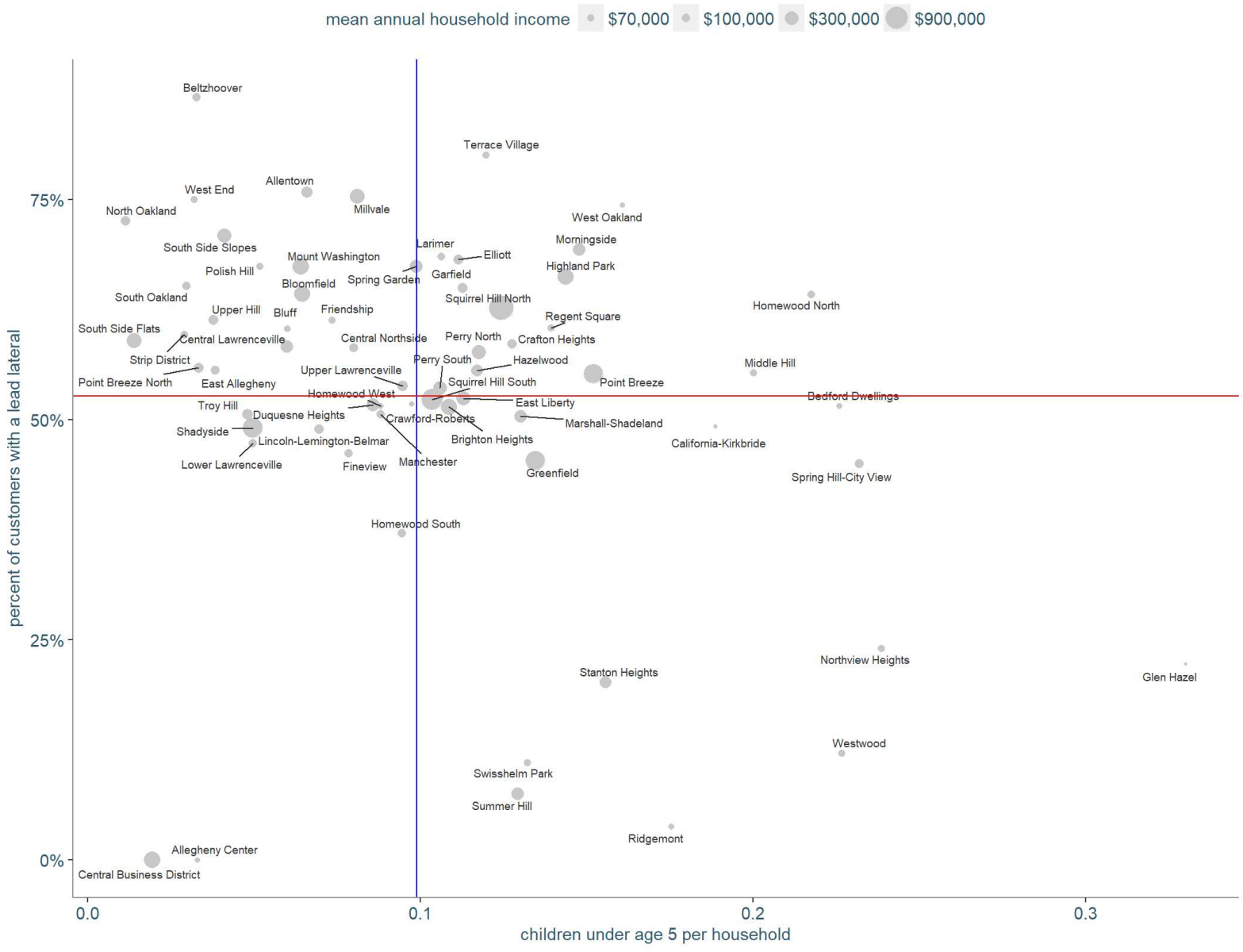

Figure 2: The incidence of children (under the age of 5 per household) given the incidence of lead laterals (percent of customers with a lead lateral) for neighborhoods in the Pittsburgh Water and Sewer Authority service area. The blue line shows the mean number of children per household. The red line shows the mean share of households with lead laterals. Uncertainty ranges are withheld on the figure for clarity but are presented in Table 3.

Figure 3 shows variation in children per household versus the mean replacement cost by neighborhood. Prioritizing lead lateral replacement in neighborhoods in the lower, right quadrant would both increase the likelihood of reducing lead exposure to children and do so in a more cost effective manner. The estimated total cost to replace all lead laterals in the neighborhoods in the lower right quadrant is $24 million (a range of $19 million to $29 million), whereas the total replacement costs in the upper left quadrant is nearly six times more at $ 150 million (a range of 120 million to $170 million) and includes 25% fewer children.

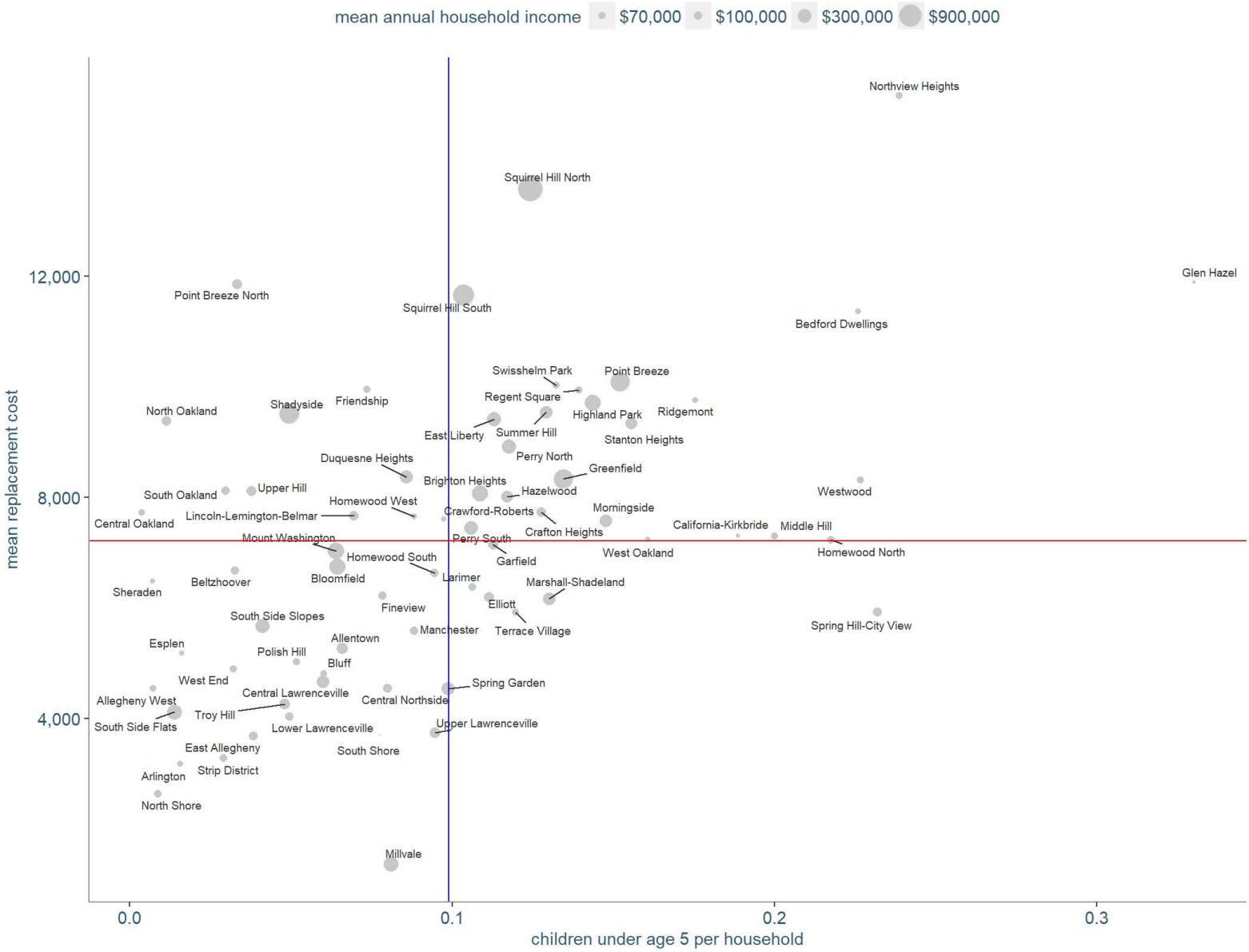

Figure 3: The incidence of children (under the age of 5 per household) given the mean cost of replacing lead laterals for neighborhoods in the Pittsburgh Water and Sewer Authority service area. The blue line shows the mean number of children per household. The red line shows the mean replacement cost. Uncertainty ranges are withheld on the figure for clarity but are presented in Table 3.

Table 3 presents summary statistics of key estimates by neighborhood. The total cost estimate to replace all lead lines is $290M (range of $190M to $490M). Table 3 suggests that cost of a lateral replacement per child protected varies from around $20,000 to over $200,000, with higher estimates mostly driven by neighborhoods with very few children under age 5. Given broad latitude over replacement strategies and limited resources for replacements, these results underscore the importance of prioritizing areas with high counts of children and lower replacement costs.

Table 3: Summary of lead inventory, children under age 5, and lead lateral replacement costs by neighborhood in the Pittsburgh Water and Sewer Authority service area. Estimates include the mean and standard deviation in parenthesis.

Neighborhood	Estimated lead laterals	Feet of lead laterals	Children under 5	Children under 5 per household	Mean replacment cost	Replacement cost per child	Total cost (millions)	Cumulative cost (millions)
Millvale	1,064	4,300 (37)	240 (37)	0.081 (0.013)	1,400 (300)	17,000 (5,000)	1.5 (0.32)	1.45 (0.32)
Spring Hill-City View	420	7,500 (170)	510 (87)	0.23 (0.041)	5,900 (1,400)	26,000 (8,400)	2.5 (06)	3.94 (0.68)
Homewood North	592	13,000 (120)	270 (49)	0.22 (0.041)	7,200 (1,400)	34,000 (11,000)	4.3 (0.81)	8.22 (1)
Westwood	21	530 (7.4)	50 (15)	0.23 (0.071)	8,300 (3,100)	40,000 (27,000)	0.17 (0.066)	8.4 (1)
Middle Hill	421	9,200 (360)	220 (63)	02 (0.057)	7,300 (1,600)	41,000 (21,000)	3.1 (0.69)	11.5 (13)
Upper Lawrenceville	619	6,900 (220)	210 (43)	0.095 (0.02)	3,700 (870)	42,000 (14,000)	2.3 (0.54)	13.8 (14)
California-Kirkbride	135	2,900 (200)	71 (23)	0.19 (0.062)	7,300 (2,800)	44,000 (30,000)	0.99 (0.37)	14.8 (15)
Spring Garden	286	3,900 (140)	220 (47)	0.099 (0.022)	4,500 (1,100)	48,000 (18,000)	1.3 (03)	16.1 (15)
West Oakland	306	6,700 (210)	120 (29)	0.16 (0.041)	7,200 (2,000)	48,000 (22,000)	2.2 (0.61)	18.3 (17)
Marshall-Shadeland	958	18,000 (250)	430 (70)	0.13 (0.022)	6,200 (1,200)	49,000 (13,000)	5.9 (11)	24.2 (2)
Terrace Village	298	5,300 (140)	190 (38)	0.12 (0.024)	5,900 (1,200)	52,000 (17,000)	1.8 (0.37)	26 (19)
Bedford Dwellings	100	3,400 (320)	260 (50)	0.23 (0.044)	11,000 (3,700)	52,000 (22,000)	1.1 (0.37)	27.1 (2)
Morningside	956	22,000 (120)	360 (58)	0.15 (0.024)	7,600 (1,400)	53,000 (14,000)	7.2 (13)	34.3 (23)

Neighborhood	Estimated lead laterals	Feet of lead laterals	Children under 5	Children under 5 per household	Mean replacment cost	Replacement cost per child	Total cost (millions)	Cumulative cost (millions)
Glen Hazel	6	210 (47)	90 (38)	0.33 (0.15)	12,000 (5,200)	58,000 (200,000)	0.071 (0.031)	34.4 (23)
Elliott	700	13,000 (160)	160 (40)	0.11 (0.029)	6,200 (1,400)	59,000 (25,000)	4.3 (0.96)	38.8 (25)
Central Northside	659	9,100 (460)	150 (34)	0.08 (0.018)	4,600 (1,100)	60,000 (23,000)	3 (0.73)	41.8 (27)
Stanton Heights	412	12,000 (150)	350 (56)	0.16 (0.025)	9,300 (2,300)	62,000 (19,000)	3.8 (0.94)	45.6 (27)
Greenfield	1,412	35,000 (200)	620 (85)	0.13 (0.019)	8,300 (1,200)	63,000 (13,000)	12 (17)	57.4 (32)
Garfield	960	21,000 (310)	220 (29)	0.11 (0.015)	7,100 (1,300)	65,000 (15,000)	6.9 (12)	64.2 (34)
Northview Heights	6	280 (31)	250 (49)	0.24 (0.048)	15,000 (6,000)	67,000 (31,000)	0.092 (0.036)	64.3 (34)
Point Breeze	1,100	33,000 (250)	380 (52)	0.15 (0.021)	10,000 (1,600)	67,000 (14,000)	11 (17)	75.4 (38)
Highland Park	1,325	38,000 (290)	470 (79)	0.14 (0.025)	9,700 (1,400)	69,000 (17,000)	13 (19)	88.3 (43)
Hazelwood	1,046	25,000 (420)	400 (68)	0.12 (0.02)	8,000 (1,200)	70,000 (18,000)	8.4 (12)	96.7 (43)
Crafton Heights	51	1,200 (22)	130 (37)	0.13 (0.037)	7,700 (2,000)	70,000 (66,000)	0.39 (01)	97.1 (44)
Manchester	404	6,700 (180)	160 (37)	0.088 (0.021)	5,600 (1,500)	71,000 (37,000)	2.3 (0.59)	99.3 (42)
Perry South	874	20,000 (230)	310 (62)	0.11 (0.021)	7,400 (1,100)	74,000 (21,000)	6.5 (0.99)	106 (43)
Larimer	401	7,700 (110)	75 (27)	0.11 (0.039)	6,400 (1,300)	74,000 (58,000)	2.6 (0.52)	108 (43)
Brighton Heights	1,438	35,000 (220)	430 (64)	0.11 (0.017)	8,100 (1,300)	76,000 (17,000)	12 (18)	120 (45)
Perry North	899	24,000 (170)	280 (43)	0.12 (0.019)	8,900 (1,400)	78,000 (18,000)	8 (13)	128 (47)

Neighborhood	Estimated lead laterals	Feet of lead laterals	Children under 5	Children under 5 per household	Mean replacment cost	Replacement cost per child	Total cost (millions)	Cumulative cost (millions)
Summer Hill	37	1,100 (81)	110 (23)	0.13 (0.027)	9,500 (3,700)	79,000 (35,000)	0.35 (0.14)	128 (46)
Homewood South	195	3,900 (110)	94 (24)	0.095 (0.025)	6,600 (1,500)	79,000 (120,000)	1.3 (03)	130 (46)
Swisshelm Park	67	2,000 (45)	83 (17)	0.13 (0.027)	10,000 (3,700)	80,000 (36,000)	0.67 (0.25)	130 (45)
Regent Square	211	6,300 (33)	63 (18)	0.14 (0.04)	9,900 (3,700)	82,000 (51,000)	2.1 (0.79)	132 (45)
Central Lawrenceville	1,127	16,000 (510)	160 (34)	0.06 (0.013)	4,700 (860)	82,000 (24,000)	5.3 (0.97)	138 (44)
Crawford-Roberts	286	6,500 (280)	120 (33)	0.097 (0.027)	7,600 (2,100)	85,000 (36,000)	2.2 (06)	140 (45)
East Liberty	577	16,000 (180)	450 (84)	0.11 (0.021)	9,400 (1,600)	86,000 (25,000)	5.4 (0.92)	145 (46)
Allentown	729	11,000 (140)	170 (47)	0.066 (0.019)	5,300 (1,200)	88,000 (40,000)	3.8 (0.89)	149 (48)
Lower Lawrenceville	417	5,200 (110)	65 (20)	0.05 (0.016)	4,000 (1,100)	93,000 (64,000)	1.7 (0.45)	151 (49)
Troy Hill	507	6,500 (110)	92 (24)	0.048 (0.012)	4,300 (1,200)	94,000 (42,000)	2.2 (06)	153 (49)
Fineview	200	3,800 (160)	84 (24)	0.078 (0.023)	6,200 (1,700)	95,000 (170,000)	1.2 (0.33)	154 (47)
Bluff	148	2,100 (140)	27 (10)	0.06 (0.022)	4,800 (1,400)	100,000 (91,000)	0.71 (02)	155 (47)
Duquesne Heights	611	15,000 (360)	230 (51)	0.086 (0.019)	8,400 (2,200)	100,000 (39,000)	5.1 (14)	160 (49)
East Allegheny	412	4,600 (150)	110 (26)	0.038 (0.0092)	3,700 (970)	100,000 (42,000)	1.5 (04)	162 (49)
Polish Hill	344	5,200 (140)	42 (9.7)	0.052 (0.012)	5,000 (1,300)	100,000 (50,000)	1.7 (0.46)	163 (5)
Bloomfield	1,779	36,000 (390)	330 (54)	0.064 (0.01)	6,700 (900)	110,000 (25,000)	12 (16)	175 (52)

Neighborhood	Estimated lead laterals	Feet of lead laterals	Children under 5	Children under 5 per household	Mean replacment cost	Replacement cost per child	Total cost (millions)	Cumulative cost (millions)
Mount Washington	2,402	50,000 (250)	380 (59)	0.064 (0.01)	7,000 (920)	110,000 (24,000)	17 (22)	192 (55)
Squirrel Hill North	1,473	60,000 (390)	470 (80)	0.12 (0.021)	14,000 (1,900)	110,000 (27,000)	20 (27)	212 (65)
Squirrel Hill South	2,017	71,000 (490)	850 (81)	01 (0.01)	12,000 (1,500)	110,000 (19,000)	24 (31)	236 (69)
Lincoln-Lemington-Belmar	818	19,000 (220)	140 (29)	0.07 (0.015)	7,700 (1,500)	110,000 (36,000)	6.3 (12)	242 (72)
Friendship	184	5,500 (34)	87 (26)	0.074 (0.022)	10,000 (2,700)	150,000 (92,000)	1.8 (0.49)	244 (71)
South Side Slopes	1,451	25,000 (350)	100 (30)	0.041 (0.012)	5,700 (820)	160,000 (78,000)	8.2 (12)	252 (74)
Strip District	43	420 (23)	17 (7.3)	0.029 (0.012)	3,300 (1,200)	160,000 (280,000)	0.14 (0.053)	252 (73)
Homewood West	167	3,800 (54)	44 (25)	0.088 (0.052)	7,700 (2,200)	190,000 (800,000)	1.3 (0.36)	253 (72)
Shadyside	1,137	33,000 (330)	420 (58)	0.05 (0.0069)	9,500 (1,100)	190,000 (35,000)	11 (12)	264 (75)
South Shore	9	100 (2.7)	1.6 (2.3)	0.078 (0.13)	3,700 (1,400)	210,000 (3,400,000)	0.033 (0.013)	264 (78)
Upper Hill	525	13,000 (140)	80 (20)	0.038 (0.0094)	8,100 (1,800)	230,000 (98,000)	4.3 (0.94)	269 (79)
West End	66	960 (37)	21 (11)	0.032 (0.017)	4,900 (1,600)	280,000 (760,000)	0.32 (0.11)	269 (78)
Beltzhoover	26	520 (52)	14 (6)	0.033 (0.014)	6,700 (1,900)	290,000 (430,000)	0.17 (0.049)	269 (77)
South Oakland	605	15,000 (450)	52 (17)	0.03 (0.0098)	8,100 (1,900)	330,000 (620,000)	4.9 (12)	274 (79)
South Side Flats	1,328	17,000 (410)	61 (20)	0.014 (0.0045)	4,100 (710)	350,000 (230,000)	5.5 (0.95)	279 (8)
Ridgemont	8	240 (0.036)	34 (16)	0.18 (0.092)	9,800 (3,700)	370,000 (8,700,000)	0.078 (0.03)	280 (82)

Neighborhood	Estimated lead laterals	Feet of lead laterals	Children under 5	Children under 5 per household	Mean replacment cost	Replacement cost per child	Total cost (millions)	Cumulative cost (millions)
Point Breeze North	334	12,000 (230)	38 (13)	0.033 (0.011)	12,000 (2,600)	440,000 (440,000)	4 (0.85)	283 (83)
Arlington	2	19 (0)	0.28 (0.2)	0.016 (0.011)	3,200 (1,200)	620,000 (4,200,000)	0.0064 (0.0024)	283 (82)
Esplen	78	1,200 (66)	1.7 (2.4)	0.016 (0.024)	5,200 (1,900)	830,000 (3,900,000)	0.4 (0.15)	284 (82)
North Oakland	175	4,900 (37)	39 (15)	0.011 (0.0042)	9,400 (2,100)	1,000,000 (1,100,000)	1.6 (0.37)	286 (79)
North Shore	3	24 (0)	1.7 (2.4)	0.0086 (0.012)	2,600 (990)	1,200,000 (15,000,000)	0.0079 (0.003)	286 (78)
Allegheny West	68	940 (48)	1.7 (2.4)	0.0072 (0.01)	4,600 (1,700)	2,100,000 (24,000,000)	0.31 (0.12)	286 (79)
Central Oakland	621	14,000 (270)	8.3 (4.9)	0.0036 (0.0021)	7,700 (1,700)	4,400,000 (14,000,000)	4.8 (11)	291 (81)
Sheraden	43	830 (25)	0.49 (0.68)	0.0069 (0.0096)	6,500 (2,400)	7,300,000 (150,000,000)	0.28 (01)	291 (76)
Allegheny Center	0	0 (0)	39 (13)	0.033 (0.011)	0 (0)	NA (NA)	0 (0)	291 (76)
Central Business District	0	0 (0)	36 (12)	0.019 (0.0065)	0 (0)	NA (NA)	0 (0)	291 (76)

4. Implications and Limitations

Municipalities and public health agencies are typically given broad discretion on how to protect their constituents from lead in drinking water. When information on the lead sources is missing, so is information describing who is at risk of exposure. As a result, municipalities and public health agencies are often unable to prioritize where to spend limited resources on prevention and mitigation.

By developing a complete lead lateral inventory and merging it with demographic data, the results presented here can help municipalities and public health agencies prioritize prevention and mitigation in neighborhoods where children are more likely to be exposed and the cost of mitigation is lower. The results demonstrate orders of magnitude in the variation of both lead laterals per child and mitigation cost per child. Strategies that do not incorporate this extreme variation - such as randomly choosing locations for interventions - are much less likely to be effective.

Whereas the lead predictions are made by address, the demographic data are only available at the block level. As a result, measures of exposure and cost effectiveness can only be prepared at the block level, meaning that interventions utilizing the results presented here will only increase the *likelihood* of protecting children at improved cost effectiveness. Any single address may not have any children, even in neighborhoods with a high number of mean children per household.

Census data in between decennial counts are notoriously uncertain, and these data dominate the uncertainty demonstrated in Table 3. In the future, improved estimated counts of children could be derived by independent surveys or possibly by utilizing the bedroom counts reported in the Allegheny County property assessment data.

While the replacement costs reflect property-specific curb setback, it should be emphasized that there appear to be no reliable sources for unit lateral replacement costs. Those assumed here align with casually understood replacement costs. As PWSA engages in replacement programs, collecting and publishing replacement cost data will improve the accuracy and precision of the replacement cost estimates.

5. Additional Resources

See UCSUR's analysis of how a lead water lateral affects property sales values in Pittsburgh (https://ucsur.pitt.edu/lead_water_laterals_2018.php).

Also, preview a prototype of "Leaducated," (http://tools.wprdc.org/lead-demo/) a web app that helps users determine if they have a lead water line.

References

Allegheny County. Department of Administrative Services, Geographic Information Systems Group. "Allegheny County Building Footprint Locations." https://data.wprdc.org/dataset/allegheny-county-building-footprint-locations (https://data.wprdc.org/dataset/allegheny-county-building-footprint-locations). Accessed Apr 17, 2018.

Allegheny County. Department of Administrative Services, Office of Property Assessments. "Allegheny County Property Assessments." https://data.wprdc.org/dataset/property-assessments (https://data.wprdc.org/dataset/property-assessments). Accessed Apr 11, 2017.

Allegheny County. Department of Real Estate. "Allegheny County Parcel Boundaries." https://data.wprdc.org/dataset/allegheny-county-parcel-boundaries (https://data.wprdc.org/dataset/allegheny-county-parcel-boundaries). Accessed Apr 11, 2017.

City of Pittsburgh. "Street Curbs." Geographic Data, 2015. http://pittsburghpa.gov/dcp/gis/gis-data-new (http://pittsburghpa.gov/dcp/gis/gis-data-new).

City of Pittsburgh. https://data.wprdc.org/dataset/pgh-neighborhoods (https://data.wprdc.org/dataset/pgh-neighborhoods). Accessed Oct 4, 2016.

Hopey, D. "PWSA will pay for public, private water line replacement." Pittsburgh Post-Gazette. Jan 29, 2018.

Pittsburgh Water and Sewer Authority. Historical lateral material information obtained through a non-disclosure agreement between the Pittsburgh Water and Sewer Authority and the University of Pittsburgh. Accessed Feb 1, 2018.

U.S. Census Bureau.https://www.census.gov/geo/maps-data/data/tiger-line.html (https://www.census.gov/geo/maps-data/data/tiger-line.html). Accessed Oct Dec 11, 2017.

US Census Bureau. 2016 American Community Survey, 5-Year Estimates. Various statistics. generated by author using American FactFinder. Retrieved from http://factfinder.census.gov (http://factfinder.census.gov). Accessed Mar 2018.