Original Research Article

ISEE Environmental Epidemiology

OPEN

Who is living near different types of US Superfund sites

A latent class analysis considering site contaminant profiles

Brittany A. Trottier[a,*], Andrew Olshan[a], Jessie K. Edwards[a], Lawrence S. Engel[a], Hazel B. Nichols[a], Alexandra J. White[b]

Background: Millions of people in the United States live near Superfund sites and may be exposed to hazardous chemicals from those sites. However, there is limited research on chemicals present at sites and the demographics of nearby communities. We aimed to identify subgroups of Superfund sites with similar contaminant profiles and evaluate whether sociodemographic characteristics vary by type of site.

Methods: We used US Environmental Protection Agency Superfund data to identify sites active in the year 2000. Census tract centroids located within 3 miles of every Superfund site were identified and a weighted average of census tract-level sociodemographics using the 2000 US Census was calculated. Superfund sites with similar contaminant profiles were identified using latent class analysis. We compared the median sociodemographic characteristics, overall and by contaminant latent class, with those of the overall 2000 US Census.

Results: We identified seven latent classes based on 12 contaminant categories from 1332 Superfund sites active in 2000. Overall, there were few differences in sociodemographics observed by the presence of any Superfund site compared with the overall US Census. After stratifying by contaminant profile, we observed evidence of disparities for two classes of sites, defined by (1) high diversity of chemical exposure and lumber industry and (2) batteries and metals, which were more likely to have higher hazard scores and to be near communities with higher proportions of non-White individuals, lower socioeconomic status, and higher social vulnerability.

Conclusion: Disadvantaged communities, with higher social vulnerability, were more likely to be near certain Superfund sites with higher hazard scores.

Keywords: Environmental exposures; Superfund Sites; Latent class analysis; Hazardous compounds

Introduction

Superfund sites are an important category of waste site and may be an especially significant point source of chemical exposure to humans. The US Environmental Protection Agency (EPA) defines a Superfund site as a contaminated site where hazardous waste has been dumped, left out in the open, or improperly managed.[1] Superfund sites can contain a single chemical or compound, but they frequently contain mixtures of compounds including, but not limited to, heavy metals, dioxins, fluorinated compounds, volatile organic compounds (VOCs), and polycyclic aromatic hydrocarbons.[1] Superfund sites are at risk of being impacted by climate change and extreme weather events, as a 2020 report projected that at least 200 sites would be at risk of extreme flooding by 2040.[2] The EPA has established the National Priority List (NPL) to determine which sites warrant further investigation and cleanup.[3] The EPA uses a mechanism called the Hazard Ranking System (HRS), a system that gives a score to each site, as a tool to assess the potential that a site may pose a threat to human health and the environment and to help determine whether a site will be placed on the NPL. The higher the score, the more potential a site has to harm human and environmental health. In 2023, there were over 1300 Superfund sites across the United States actively emitting contaminants.[1]

It has been estimated that more than 73 million people live within 3 miles of a Superfund site and almost 21 million people live within 1 mile of a site.[4] Several studies have found that residential proximity to Superfund sites is associated with the occurrence of adverse human health effects, ranging from poor birth outcomes to chronic diseases.[5–8] Studies have also detected

[a]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; and [b]National Institute of Environmental Health Sciences, Durham, North Carolina

Data and code are available upon reasonable request to the first author.

This work was funded in part by the intramural program of the National Institute of Environmental Health Sciences, Z01-ES044005, Z1AES103332.

SDC *Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.environepidem.com).*

**Corresponding Author. Address: Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, 170 Rosenau Hall, 135 Dauer Drive, Chapel Hill, NC 27599. E-mail: brittany.trottier@nih.gov (B.A. Trottier).*

Copyright © 2025 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of The Environmental Epidemiology. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Environmental Epidemiology (2025) 9:e363

Received 29 August, 2024; Accepted 10 December, 2024

Published online 3 February 2025

DOI: 10.1097/EE9.0000000000000363

What this study adds:
Previous studies that assessed socioeconomic characteristics of populations near Superfund sites focused on a single site or single state. This study expands on this by capturing Superfund sites nationwide and considering the types of contaminants present at the sites and industries contributing to the sites. This study also included additional dimensions of the socioeconomic factors of census tracts such as social vulnerability.

contaminants from Superfund sites in the air, soil, and dust of nearby homes.[9,10] We previously reported that living in closer proximity to metal-containing Superfund sites was associated with higher toenail lead concentrations, particularly for non-Hispanic Black women.[11] These findings were consistent with studies that have found that race correlates with both environmental exposures and social stressors, suggesting the potential for multiple colocated sources of exposure or enhanced susceptibility to contaminant exposure.[12] However, to date, there has been little done to comprehensively characterize the sociodemographic factors related to living near Superfund sites using a national lens. Additionally, given the wide range of chemicals that may be contaminating these sites, a better characterization of the types of pollutant mixtures at each site may help inform which sites may be more hazardous to neighboring communities.

The aim of our study was to classify US Superfund hazardous waste sites into subgroups based on the contaminant profiles and to evaluate how these site profiles were related to both neighborhood and individual sociodemographic characteristics using data from the US Census.

Materials and methods

Datasets

We used a dataset from the US EPA[13] that contained information for all Superfund sites, including site location, industry, presence of chemical(s) emitted, hazard ranking score, and the contaminated media surrounding the site. We limited the dataset to those sites that were considered active (i.e., actively releasing chemicals into the environment) in the year 2000, as determined by the NPL.

For sociodemographic characteristics of people who live near Superfund sites, we obtained data from the 2000 US Census, which matched the year of the Superfund site data. The 2000 Census was chosen because it had the most complete information for the factors of interest. The proportion of the population with a given census tract characteristic was estimated by the median percentage of the population with that characteristic and included the following: race (percent non-Hispanic White; percent non-Hispanic Black; Hispanic; percent non-Hispanic Asian, native Hawaiian, or other Pacific Islander; non-Hispanic American Indian, Alaskan native, Eskimo, or Aleut), household income (percent less than \$50,000, \$50,000–\$99,999, \$100,000+), educational attainment (percent with no high school diploma; high school diploma; some college or associate's degree; bachelor's degree or higher), housing type (percent of single-family home or duplex; small multiunit complex, 3–9 units; large multiunit complex, 10 or more units; mobile home or other), and Social Vulnerability Index (SVI) overall percentile ranking. The SVI data were obtained from the Centers for Disease Control and Prevention Agency for Toxic Substances and Disease Registry database for the 2000 release.[14] The SVI uses 16 US Census measures (e.g., employment status, health insurance status, housing cost burden, household characteristics, housing type, access to vehicle, race, and education) to characterize the social vulnerability of every census tract in the United States, with a higher SVI indicating more vulnerability and less ability to adapt to hazards and other community-level stressors. The SVI is a commonly used composite measure for environmental exposures and climate change.

We linked the US EPA Superfund data to the 2000 US Census sociodemographic characteristics at the census tract level. We considered the census tract population to be exposed if there was a Superfund site within 3 miles of the tract centroid. To capture the overall demographic profile of the population living within 3 miles of each Superfund site, we identified all census tract centroids located within 3 miles of every Superfund site in the study and generated a weighted average of the census information from those centroids, effectively creating a representative demographic profile of the population located around each Superfund site.[15] This weighted average creates a more accurate summary of the demographics of the populations within the buffer region rather as these buffers may cross into multiple census tracts. It better captures the population potentially affected by contamination from a Superfund site than does a single census tract containing site. We evaluated weighted average medians as the weighted averages were not normally distributed.

Statistical analysis

We used latent class analysis (LCA) to identify subgroups of Superfund sites with similar contaminant profiles. LCA is a dimension-reduction technique that creates subgroups (classes) using a probabilistic approach to uncover clusters in data by grouping subjects using prespecified features based on unobservable membership.[16]

In this application of LCA, we used the Superfund contaminant indicator variables as the inputs to determine class membership. Specifically, these defining variables included the following dichotomous chemical classifications used by Agency for Toxic Substances and Disease Registry for contaminants found at the Superfund sites: metals (yes/no), benzidines/aromatic amines (yes/no), dioxins/furans/PCBs (yes/no), halogenated compounds (yes/no), inorganic substances (yes/no), nitrosamines/ethers/alcohols (yes/no), organophosphates (yes/no), phenols/phenoxy acids (yes/no), phthalates (yes/no), radionuclides (yes/no), and VOCs (yes/no).

We evaluated different numbers of classes for the LCA, ranging from four to eight, to identify the best-fitting model. Bayesian information criteria (BIC), Akaike information criteria (AIC), maximum log-likelihood, entropy, and interpretability were used to choose the final model. However, the BIC was prioritized over the other statistical tests for final model selection because it is a well-established measure of goodness of fit of a statistical model.[16] Superfund sites were then assigned to the class for which they had the highest probability of having similar contaminant profiles.

Superfund site latent classes were described based on their prevalence, probability of including certain contaminants, and their most common industries. The industry of site variable was further subdivided into another variable, "subindustry," which included more detailed information about the activities at the site (e.g., metal processing plant, industrial waste facility, and lumber processing plant) and was included in the descriptive table if more than 10% of the sites fell within the subindustry category. Latent classes were named using the chemical profiles identified and information from the industry subcategory. A class was named a "high diversity exposure" class if more than half of the 12 defined chemical categories were present with high probabilities of being included in the latent class and was named "low diversity exposure" if there were fewer than 6 chemical categories represented.

We obtained the posterior probability (the average probability of the model accurately predicting class membership for each site) and estimated the mean and median HRS score (a score that determines threat to human and environmental health) for each of the classes.

We determined the median Census characteristics for the census tracts with centroids within 3 miles of a Superfund site and compared these to the overall US population medians (i.e., overall 2000 US Census medians). We also considered how these census characteristics varied across Superfund classes.

We conducted a sensitivity analysis considering all census tract centroids located within 1 mile, rather than 3 miles, of a Superfund site. The 1- and 3-mile distances are consistent with EPA reports characterizing populations living near Superfund sites.[17]

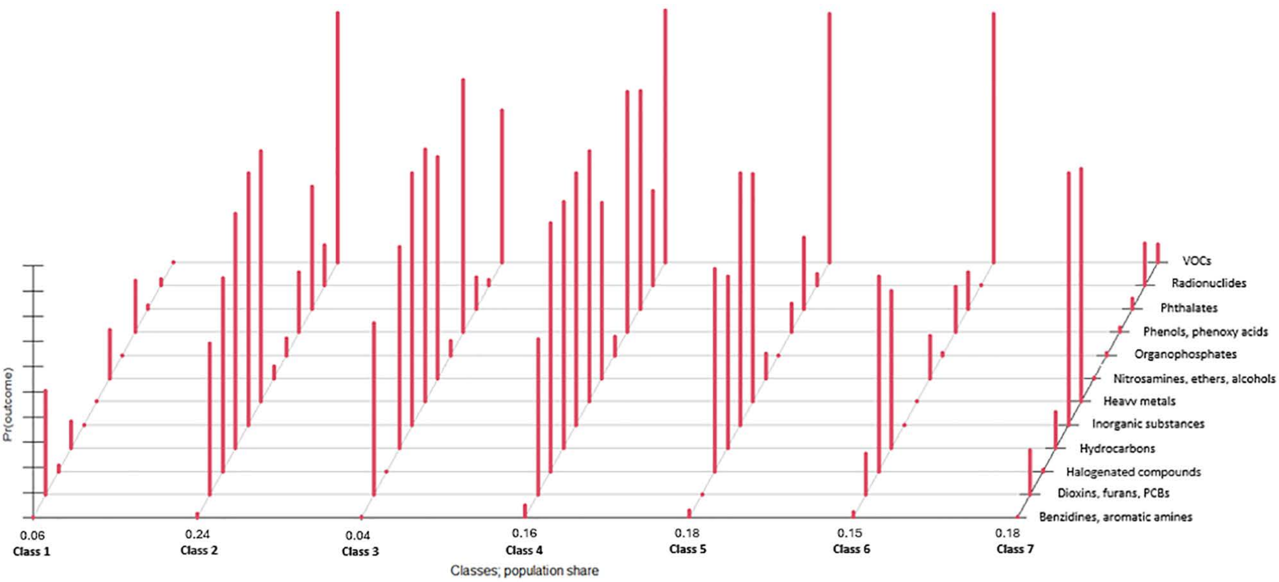

Figure 1. Distribution of chemical categories across latent classes.

Table 1.

Probability of distribution of contaminant indicators in each latent class

	Class 1 (84 sites)	Class 2 (299 sites)	Class 3 (47 sites)	Class 4 (236 sites)	Class 5 (227 sites)	Class 6 (195 sites)	Class 7 (244 sites)
Benzidines, aromatic amines	0.00	0.01	0.00	0.05	0.03	0.02	0.00
Dioxins, furans, polychlorinated biphenyls	0.41	0.60	0.68	0.61	0.00	0.16	0.18
Halogenated hydrocarbons, halogenated volatile organic compounds	0.02	0.77	0.00	0.99	0.80	0.77	0.01
Hydrocarbons	0.11	0.93	0.80	0.98	0.68	0.62	0.14
Inorganics	0.00	1.00	1.00	1.00	1.00	0.00	1.00
Metals	0.00	1.00	1.00	0.99	0.90	0.00	0.92
Nitrosamines, ethers, alcohols	0.19	0.05	0.88	0.70	0.10	0.17	0.00
Organophosphates	0.00	0.07	0.06	0.07	0.00	0.01	0.01
Phenols, phenoxy acids	0.20	0.24	1.00	0.95	0.11	0.18	0.02
Phthalates	0.01	0.49	0.12	0.86	0.28	0.14	0.04
Radionuclides	0.02	0.16	0.02	0.38	0.05	0.00	0.17
Volatile organic compounds	0.00	0.99	0.60	1.00	0.99	0.99	0.07

LCA was performed using the poLCA package in R Studio version 4.3.1. Descriptive analyses were performed in SAS version 9.4.

Results

Latent class analysis of Superfund sites

We identified 1332 Superfund sites that were actively releasing chemicals in the year 2000. Using LCA, we determined that the seven-class model best fits the data (model fit statistics can be found in Supplemental Table 1; http://links.lww.com/EE/A325). Although the eight-class model had slightly better AIC and likelihood outcomes, the seven-class model was ultimately chosen due to the lower BIC, highest posterior probability, and clearer interpretability of the model compared with the eight-class model.

Characteristics of Superfund site classes

The latent classes were characterized by varying distributions of chemicals (Figure 1 and Table 1). Class 3 contained the fewest number of Superfund sites (47 sites) while class 2 had the most (299 sites) (Supplemental Table 2; http://links.lww.com/EE/A325). Classes 2–5 had exposure to a higher number of chemical categories compared with the other classes. Class 7 had the highest median HRS score (HRS) while class 6 had the lowest HRS score (42.2 vs. 36.4) (Table 2).

The classes were named based on their contaminant exposure and industries as follows (Supplemental Table 3; http://links.lww.com/EE/A325, footnotes of Table 2): Class 1 was named "low diversity exposure, lumber and wood industry" because there were fewer chemical categories included in this class than other classes and lumber and wood products was the most common industry subcategory (19% of sites); class 2 was named "high diversity exposure, industrial and municipal landfills"; class 3 was almost solely wood and lumber (64% of sites) and had a large number of chemical categories included so it was named "high diversity exposure, lumber and wood industries"; class 4 had sites distributed across waste, landfill, and chemical industries, thus was named "high diversity exposure, landfills and industrial waste facilities, chemical industries"; class 5 also had several chemicals in the class and was named "high diversity exposure, landfills and waste, metal and electronic industries"; class 6 was named "manufacturing and processing facilities, electronic/electrical and chemical industries"; class 7 was named "batteries and scrap metal industries, metal processing."

Characteristics of census tracts within 3 miles of Superfund sites

The overall median values for race, income, education, housing, and SVI of the census tracts within 3 miles of any Superfund

Table 2.

Sociodemographic characteristics of US Census tracts within 3 miles of any Superfund site compared with overall US Census characteristics (2000 Census year) and with Superfund sites characterized in latent classes based on contaminant profiles

	2000 US Census	Any Superfund Site within 3 miles (1332 sites)	Superfund site within 3 miles, by Superfund contaminant latent classes[a]						
			Class 1 (84 sites)	Class 2 (299 sites)	Class 3 (47 sites)	Class 4 (236 sites)	Class 5 (227 sites)	Class 6 (195 sites)	Class 7 (244 sites)
	Median	Median	Median	Median	Median	Median	Median	Median	Median
Hazard ranking score		40.1	36.8	40.4	41.9	41.0	40.7	36.4	42.2
Race									
Percent population non-Hispanic, White	80.8	84.1	83.2	83.4	76.4	85.6	86.2	86.5	78.4
Percent population non-Hispanic, Black	2.7	3.0	2.5	4.1	5.7	3.9	2.8	2.1	2.7
Percent population Hispanic origin (any race)	3.6	3.4	3.6	3.3	2.9	3.2	3.4	2.9	4.8
Percent population non-Hispanic, Asian, native Hawaiian, or other Pacific Islander	0.9	0.8	0.8	0.9	0.8	1.0	0.7	0.8	0.7
Percent population non-Hispanic, American Indian, Alaskan native, Eskimo, Aleut	1.7	0.9	0.9	0.9	1.3	0.8	0.9	0.8	0.9
Household income									
Percent of population with <$50,000 household income	62.7	61.5	60.9	60.8	70.0	59.1	62.7	56.0	65.7
Percent of population with $50,000 to $99,999 household income	28.6	30.0	30.0	30.9	25.8	31.8	28.9	33.0	27.2
Percent of population with $100,000 or more household income	7.0	7.8	7.0	8.2	6.3	8.9	7.3	9.6	6.6
Education (persons age 25 or older)									
Percent of population with no high school diploma	18.0	17.3	17.4	17.2	18.6	17.4	17.9	15.5	19.6
Percent of population with high school diploma	29.6	31.8	31.7	32.2	32.4	31.8	32.8	31.0	30.8
Percent of population with some college or associate's degree	26.8	26.8	28.3	26.3	28.3	26.3	26.7	27.3	26.9
Percent of population with bachelor's degree or higher	17.8	18.6	18.5	19.2	16.6	19.7	17.4	20.6	16.7
Housing									
Percent of population in single-family homes or duplex	78.0	78.4	77.3	79.6	76.1	78.4	78.2	79.3	77.2
Percent of population in small multiunit complex, 3–9 units	5.4	7.3	6.9	7.8	8.6	7.5	7.2	6.9	6.3
Percent of population in large multiunit complex, 10 or more units	3.5	4.4	3.8	5.2	5.5	4.4	3.7	4.7	3.8
Percent of population in mobile home or other type of housing	0.7	4.9	6.0	4.0	7.1	4.1	5.0	5.0	7.0
SVI									
SVI: percentile of the proportion minority	0.5	0.4	0.4	0.4	0.5	0.4	0.4	0.3	0.5
SVI: percentile of the proportion of persons below the poverty estimate	0.5	0.4	0.5	0.4	0.6	0.4	0.5	0.4	0.5
SVI: overall percentile ranking	0.5	0.5	0.5	0.4	0.7	0.4	0.5	0.4	0.5

[a]Class names: class 1 (low diversity exposure, lumber and wood), class 2 (high diversity exposure, industrial/municipal landfills), class 3 (high diversity exposure, lumber and wood), class 4 (high diversity exposure, landfills, industrial waste, chemical industries), class 5 (high diversity exposure, landfills/waste, metal/electronic industries), class 6 (manufacturing/processing, electric, chemicals), class 7 (batteries, scrap metals, metal processing).

site were generally similar to the general US population characteristics with the exception of the percentage of the population residing in mobile homes; however, differences emerged when stratified by the seven Superfund latent classes that were identified (Table 2).

Race

Class 3 (high diversity exposure, lumber and wood) had the lowest proportion of the population that was non-Hispanic White (76.4% vs. 80.8% in the general population), the highest proportion of non-Hispanic Black (5.7 vs. 2.7 in the general population), and the highest proportion American Indians/Alaska Native among the classes; however it was lower than the census median (1.3 vs. 1.7 in the general population) (Table 2). Class 7 (batteries and metals) had the highest proportion of Hispanic origin (4.8) compared with the other classes and the lowest proportion of Asian, Native Hawaiian, or other Pacific Islander (0.7). The lowest proportion of non-Hispanic Black individuals was in class 6 (2.1), which was the class that also had the largest proportion being non-Hispanic White (86.5), lowest proportion of Hispanic origin (2.9), and lowest proportion of American Indian, Alaskan native, Eskimo or Aleut (0.8). Class 4 had the highest proportion of Asian, native Hawaiian, or other Pacific Islanders (1.0). Classes 1 (low diversity exposure, lumber) and 2 (industrial/municipal landfills) followed similar patterns for general census race demographics.

Income

Class 3 (high diversity exposure, lumber and wood) had the lowest proportion of high-income earners (6.3) while class 6 (manufacturing/processing, electric, chemicals) had the highest proportion of high-income earners (household income of $100,000 or more) compared with the other classes and general US Census income (9.6 vs. 7.0 in the general population). Class 3 also had the highest proportion in the lowest income tier (income < $50,000) while class 6 had the lowest proportion in the lowest income tier (70 and 56). Classes 3 and 7 also had higher proportions in the lowest income tiers when compared with the Census (70 and 65.7 vs. 62.7). The other classes followed similar patterns compared with the overall 2000 US Census characteristics.

Education

Class 3 (high diversity exposure, lumber and wood) and class 7 (batteries, scrap metal, metal processing) consisted of the least educated populations with the highest proportion with no high school diploma, which was higher than the general population educational attainment (18.6 and 19.6 compared with 18.0 in the general population); class 6 was the most educated when comparing the education categories (20.6).

Housing type

Housing type was distributed similarly across the latent classes. Of note, class 3 had the highest proportion of living in mobile homes or other types of housing compared with other classes and the overall median of the classes (7.1 for class 3 compared with an overall median of 4.9). Additionally, class 6 had the highest proportion living in a single-family home or duplex compared with other classes and the overall median of the classes (79.3 for class 6 compared with an overall median of 78.4).

Social Vulnerability Index

A high SVI percentile indicated higher vulnerability. The SVI measures for populations near Superfund sites were generally slightly lower than those of the 2000 Census population. The highest overall SVI percentile ranking was in class 3 (0.70); classes 4 and 6 each had the lowest overall SVI percentile ranking (0.40). A high SVI percentile indicated higher vulnerability. Class 3 also had the highest percentile of proportion minority and persons below poverty compared to the other classes (0.5 and 0.6, respectively).

The overall results from the sensitivity analysis using a 1-mile distance showed similar patterns by sociodemographic characteristics and latent class (Supplemental Table 4; http://links.lww.com/EE/A325).

Discussion

In this comprehensive cross-sectional study of active US Superfund sites, we observed seven latent classes that explained most of the variability in the chemical contaminants at each site. While the sociodemographic characteristics of individuals living within 3 miles of a Superfund site were similar to those of the US general population in 2000, exposure disparities by race, income, and education were observed when considering different types of Superfund sites, as defined by contaminant profiles and industry categories. Understanding the demographics of communities near the different types of Superfund sites is important for public health practitioners to develop risk messaging and mitigation strategies because the communities will view and understand risk differently and have different abilities to respond (e.g., based on income).[18]

A common theme emerged among classes 3 and 7 when comparing race, income, educational attainment, and SVI to the general population and the other latent classes. When comparing the racial and ethnic distribution of the census tracts, class 3 (high diversity exposure, lumber and wood) and class 7 (batteries, scrap metal, metal processing) were comprised of more minoritized populations compared with the general population and when compared with the other latent classes. The populations that were exposed to Superfund sites in classes 3 and 7 generally had lower income levels and lower educational attainment. This is important as Classes 3 and 7 also had the highest HRS score, which suggests that these sites have a greater potential for threat to human health.

These communities exposed to Superfund sites in classes 3 and 7 also had higher median social vulnerability rankings, indicating the communities exposed to these Superfund site classes have a reduced ability to respond to external stresses on human health compared with other populations. Their higher SVI rankings capture pre-existing socioeconomic factors that make these communities more vulnerable and less able to adapt and respond to changes in environment, disasters, and social circumstances.[19] Their higher SVI indicates that communities living near these classes of Superfund sites have fewer resources to cope with potential exposures from the sites than communities with more resources. It is also plausible that these communities are living near several colocated sources of hazardous exposures, not just Superfund sites, though we lacked the data to assess this. Achieving a better understanding of the communities that are disproportionately experiencing these environmental injustices is important for targeted interventions and will be increasingly necessary with climate change. An US Government Accountability Office report found that at least 60% of Superfund sites are highly vulnerable to impacts of climate change,[20] posing an even larger threat to communities that lack the resources to adequately respond.

A 2024 study found that all socioeconomic indicators of disadvantage were associated with a higher burden of exposure to carcinogenic air emissions from industrial sites, using data from the Toxic Release Inventory.[21] For example, the authors reported that there were 18% higher odds of living in a tract with the highest air emissions for every 10% increase in the proportion of African Americans in the tract. Additionally, they found that persistent poverty was associated with greater odds of having the highest air emissions. Given the colocation of Superfund sites with industrial emissions,[22] these findings in conjunction with what we observed here suggest that disadvantaged communities may be experiencing an inequitable burden of exposure to multiple sources of pollutants. Further, a 2015 study that assessed whether community socioeconomic disadvantage followed the siting of environmental hazards or vice versa found that hazardous waste facilities were more often placed in areas with already disproportionate and increasing concentrations of people of color.[23]

Conversely, the population living in proximity to Superfund sites in class 4 (landfill, waste, chemicals) and class 6 (manufacturing/processing, electric, chemicals) were more likely to have the highest educational attainment and the highest income compared with the other Superfund site classes. These sites also had some of the lowest site median HRS scores, indicating that these are not the most hazardous sites on the NPL. Populations exposed to Superfund sites in class 6 also had the lowest SVI rankings, which suggests that the tracts are more likely to have the social support and resources to respond to potential exposure from sites.

In 2020, using information from the 2015–2018 American Community Survey and 2019 Brownfields data, the EPA reported that higher proportions of those identifying as "Black," "Asian," "Hawaiian/Pacific Islander," and "other" live within 3 miles of a Superfund or Brownfield site compared with the US average (e.g., 14.8% of population living within 3 miles of a Superfund site are Black, compared with 12.7% of total US population).[24] Additionally, the EPA report shows that 15.5% of the population living within 3 miles of a Superfund/Brownfield site have incomes below the poverty level compared with 13.7% of the US population. While the EPA reports are also based on the US population in 2020, discrepancies between our findings (e.g., our non-Hispanic Black population median estimate of 2.66% vs. EPA reporting a mean estimate of 12.6% non-Hispanic Blacks) may be due in part to the differences in timing of the studies (US population in 2020 vs. US population in 2000 for our study) as well as the inclusion of the Brownfields, which are more likely to be in urban areas and populations surrounding Brownfields include more minority, low income, linguistically isolated, and are less likely to have a high school education than the US population as a whole.[24] Additionally, non-Hispanic black populations are highly concentrated and comprise 50% or more of the total population for 96 counties but accounted for less than 6% of the population in 3141 counties, which may explain the difference between our estimates and the estimates reported by the EPA.[25] The EPA report did not explore how the characteristics of the site itself were related to neighboring population characteristics.

Our findings are also similar to a study in Rhode Island, which found that more hazardous Superfund sites were clustered in neighborhoods with lower household income levels and larger percent Black populations, consistent with the greater environmental health burdens often placed on lower-income and minoritized communities.[22,26] Similarly, a study of Superfund sites in Illinois reported that areas within a 5-mile radius of a site had significantly higher racial minority percentages compared with areas outside the radius. While the study did not find that income disparities were statistically significant, they did report that average household income was higher outside 1-, 2-, and 5-mile radii compared with within the radii.[27] Other studies of environmental injustice have also found that lower-income, minoritized, and disenfranchised populations have inequitable exposure to environmental contamination that contributes to poorer health outcomes.[28]

Our study improves upon previous related research and EPA reports by including additional factors reflecting socioeconomic status (SES) and vulnerability that have not previously been examined. For example, the SVI provides a more holistic measure of community-level SES and related susceptibility to external stressors, given that it is measured at the census block group level and includes neighborhood-level indicators of poverty, education, housing, and employment. Further, we characterized the sites by their contaminant profiles, which is when we saw the differences in population demographics near sites truly emerge. This is an important strength of this study as it helps to better characterize these heterogeneous Superfund sites to better represent real-world exposure patterns.

A limitation of this study is that it does not fully account for the fate and transport of the chemicals that are released from Superfund sites; as such, individuals within the population living within 3 miles of the site may be differentially exposed to each of the chemicals listed at the site depending on the movement of each specific chemical or compound. We focused on individuals living within 1 and 3 miles of a site, as they are more likely than individuals living further away, to experience exposure to contaminants, although these exposures likely vary by compound and how it is released into the environment. Given the mixture of exposures at each site, individual transport dynamics for certain chemicals may not adequately capture the population-level exposure. The choice of the 3-mile buffer may also obscure differences in exposure across communities within that buffer by distance from the Superfund site; however, our sensitivity analysis using a 1-mile buffer demonstrated similar results. Further, we did not have the necessary information to evaluate the timing of any remediation processes and whether or not that impacted our findings. Another limitation of the study is that the Census data relies on self-reporting, which can introduce bias. Additionally, the US Census is known to undercount racial and ethnic minorities,[29] so our results may underrepresent the racial differences in those who live near Superfund sites.

In this US-wide study of Superfund sites, we identified distinct subgroups of sites with different contaminant profiles and observed that individuals living closest to the most hazardous sites with the highest number of categories of chemicals were more likely to be from disadvantaged communities with a higher social vulnerability. These findings are important as Superfund sites may be an important point source of exposure to hazardous chemicals. Superfund sites are added to the NPL faster than they can be removed due to the complexity and cost of cleanup of the various contaminants. Most people do not know they live near these hazardous sites and are exposed to chemicals that may be harmful to their health. It is important for future studies to examine the current and potential health impacts on nearby communities of the different types of Superfund sites and their toxicant profiles for improved risk communication and to prevent or mitigate harmful exposures.

Conflicts of interest statement

The authors declare that they have no conflicts of interest with regard to the content of this report.

References

1. EPA. What is Superfund? US Environmental Protection Agency. Available at: https://www.epa.gov/superfund/what-superfund. Accessed 5 August 2020.
2. Scientists UoC. A Toxic Relationship: Extreme Coastal Flooding and Superfund Sites. 2020. Available at: http://www.ucsusa.org/resources/toxic-relationship. Accessed 1 September 2021.
3. EPA. Superfund: CERCLA Overview. US Environmental Protection Agency. Available at: https://www.epa.gov/superfund/superfund-cercla-overview#:~:text=The%20Comprehensive%20Environmental%20Response%2C%20Compensation%2C%20and%20Liability%20Act%20of%201980,waste%20at%20these%20sites%3B%20and. Accessed 25 September 2020.
4. EPA. Population Surrounding 1,857 Superfund Remedial Sites. United States Environmental Protection Agency. Available at: https://www.epa.gov/sites/production/files/2015-09/documents/webpopulationrsuperfundsites9.28.15.pdf. Accessed 1 January 2022.
5. Little BB, Reilly R, Walsh B, Vu GT. Cadmium is associated with type 2 diabetes in a Superfund site lead smelter community in Dallas, Texas. *Int J Environ Res Public Health*. 2020;17:4558.
6. Brender JD, Maantay JA, Chakraborty J. Residential proximity to environmental hazards and adverse health outcomes. *Am J Public Health*. 2011;101:S37–S52.
7. Sergeev AV, Carpenter DO. Hospitalization rates for coronary heart disease in relation to residence near areas contaminated with persistent organic pollutants and other pollutants. *Environ Health Perspect*. 2005;113:756–761.
8. Suter MA, Aagaard KM, Coarfa C, et al. Association between elevated placental polycyclic aromatic hydrocarbons (PAHs) and

PAH-DNA adducts from Superfund sites in Harris County, and increased risk of preterm birth (PTB). *Biochem Biophys Res Commun*. 2019;516:344–349.
9. Loh MM, Sugeng A, Lothrop N, et al. Multimedia exposures to arsenic and lead for children near an inactive mine tailings and smelter site. *Environ Res*. 2016;146:331–339.
10. Zota AR, Schaider LA, Ettinger AS, Wright RO, Shine JP, Spengler JD. Metal sources and exposures in the homes of young children living near a mining-impacted Superfund site. *J Expo Sci Environ Epidemiol*. 2011;21:495–505.
11. Trottier BA, Niehoff NM, Keil AP, et al. Residential proximity to metal-containing Superfund sites and their potential as a source of disparities in metal exposure among U.S. women. *Environ Health Perspect*. 2023;131:37701.
12. Morello-Frosch R, Lopez R. The riskscape and the color line: examining the role of segregation in environmental health disparities. *Environ Res*. 2006;102:181–196.
13. Environmental Protection Agency. Superfund Enterprise Management System. Available at: https://cumulis.epa.gov/supercpad/cursites/srchsites.cfm. Accessed 10 September 2022.
14. Centers for Disease Control and Prevention. CDC/ATSDR Social Vulnerability Index. Available at: https://www.atsdr.cdc.gov/place-health/php/svi/svi-data-documentation-download.html. Accessed 1 July 2020.
15. Quist AJL, Holcomb DA, Fliss MD, Delamater PL, Richardson DB, Engel LS. Exposure to industrial hog operations and gastrointestinal illness in North Carolina, USA. *Sci Total Environ*. 2022;830:154823.
16. Naldi L, Cazzaniga S. Research techniques made simple: latent class analysis. *J Invest Dermatol*. 2020;140:1676–1680.e1.
17. Environmental Protection Agency. Population Surrounding 1,881 Superfund Sites. 2024. Updated July 2023. Available at: https://www.epa.gov/system/files/documents/2023-08/FY22%20Population%20Estimates%20Superfund%20Final.pdf. Accessed 31 July 2024.
18. ITRC (Interstate Technology Regulatory Council). Risk Communication Toolkit RCT-1. Available at: https://rct-1.itrcweb.org/. Accessed 1 June 2024.
19. Mah JC, Penwarden JL, Pott H, Theou O, Andrew MK. Social vulnerability indices: a scoping review. *BMC Public Health*. 2023;23:1253.
20. U.S Government Accountability Office. Superfund: EPA Should Take Additional Actions to Manage Risks from Climate Change Effects. 2021. Available at: https://www.gao.gov/products/gao-21-555t#:~:text=Federal%20data%20suggests%20about%2060,manage%20risks%20at%20these%20sites. Accessed 1 August 2024.
21. Madrigal JM, Flory A, Fisher JA, et al. Sociodemographic inequities in the burden of carcinogenic industrial air emissions in the United States. *JNCI*. 2024;116:737–744.
22. Mohai P, Saha R. Which came first, people or pollution? Assessing the disparate siting and post-siting demographic change hypotheses of environmental injustice. *Environ Res Lett*. 2015;10:115008.
23. Paul Mohai RS. Which came first, people or pollution? Assessing the disparate siting and post-siting demographic change hypotheses of environmental injustice. *Environ Res Lett*. 2015;10. doi:10.1088/1748-9326/10/11/115008.
24. Environmental Protection Agency. Population Surrounding 27,030 Brownfields Sites. https://www.epa.gov/sites/default/files/2015-09/documents/webpopulationbrownfieldsites.9.28.15.pdf. Accessed 28 June 2024.
25. Jesse McKinnon. Census 2000 Brief: The Black Population: 2000. Economics and Statistics Administration/U.S. Department of Commerce; 2001. https://www.census.gov/library/publications/2001/dec/c2kbr01-05.html. Accessed 1 June 2024.
26. Pascual KJ, Palosaari A, Ochoa J, Dreyer C. Environmental health burdens and socioeconomic status in Rhode Island: using geographic information systems to examine health disparities in medical school. *Cureus*. 2020;12:e9816.
27. Maranville AR, Ting TF, Zhang Y. An environmental justice analysis: Superfund sites and surrounding communities in Illinois. *Environ Justice*. 2009;2:49–58.
28. Nixon R. Slow Violence, gender, and the environmentalism of the poor. In: *Slow Violence and the Environmentalism of the Poor*. Harvard University Press; 2011:128–149.
29. Deborah Stempowski. Counting Every Voice: Understanding Hard-to-Count and Historically Undercounted Populations. Available at: https://www.census.gov/newsroom/blogs/random-samplings/2023/10/understanding-undercounted-populations.html#:~:text=For%20example%2C%20according%20to%20the,Native%20population%20living%20on%20reservations. Accessed 28 December 2023.