International Journal of
Environmental Research and Public Health

MDPI

Article

Environmental Contamination and Early-Onset Colorectal Cancer: A Statewide Ecological Analysis of Superfund Sites in Texas

Lihchyun Joseph Su [1,*], Emina H. Huang [2], Sean G. Young [1], Ping-Ching Hsu [3], MinJae Lee [4], Heather Kitzman [5], Michael Bauer [6], Joyane Eriom [1], Kazmi Syed [2] and Shuk-Mei Ho [6]

1 O'Donnell School of Public Health, UT Southwestern Medical Center, Dallas, TX 75390, USA; sean.young@utsouthwestern.edu (S.G.Y.); joyane07@gmail.com (J.E.)
2 School of Medicine, UT Southwestern Medical Center, Dallas, TX 75390, USA; emina.huang@utsouthwestern.edu (E.H.H.)
3 Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA
4 McGovern Medical School, UT Health Houston, Houston, TX 77030, USA
5 Parkland Health, Dallas, TX 75235, USA
6 School of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA
* Correspondence: lihchyun.su@utsouthwestern.edu

Highlights

Public health relevance—How does this work relate to a public health issue?

- County-level environmental contamination from Superfund sites may be associated with the rising burden of early-onset colorectal cancer, highlighting environmental exposures as a hypothesis for further investigation.
- Counties with greater socioeconomic disadvantages and larger Black populations may also experience disproportionate environmental burdens. These overlapping county-level patterns warrant environmental justice research using individual-level exposure and outcome data.

Public health significance—Why is this work of significance to public health?

- This exploratory ecological study identifies a county-level association between Superfund site presence and the availability of a reportable early-onset colorectal cancer incidence estimate; it does not establish individual exposure or causation.
- The findings identify environmental and demographic patterns that can guide more spatially resolved, individual-level studies of early-onset colorectal cancer.

Public health implications—What are the key implications or messages for practitioners, policy makers and/or researchers in public health?

- Public health agencies and researchers should consider enhanced environmental monitoring and individual-level studies in environmentally burdened communities before using these findings to guide clinical screening policies.
- Future research should incorporate geocoded residential histories, distance and exposure-pathway measures, site characteristics and remediation status, individual risk factors, and spatial statistical methods.

Check for updates

Academic Editors: Kai Yang and Beibei Wang

Received: 21 August 2026
Revised: 30 October 2026
Accepted: 12 November 2026
Published: 24 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Abstract

Superfund sites are an important public health concern in Texas and may disproportionately affect underserved communities. This exploratory ecological study examined whether county-level Superfund site presence was associated with having a publicly reportable early-onset colorectal cancer (EOCRC) incidence estimate across all 254 Texas counties.

Superfund site data were obtained from the Texas Commission on Environmental Quality, demographic data from the U.S. Census Bureau, and cancer incidence data from the Texas Cancer Registry for 2018–2022. Because county estimates based on small numbers were suppressed, the outcome distinguished counties with a non-zero, non-suppressed EOCRC estimate from counties with a zero or suppressed estimate; it should not be interpreted as a clinical threshold for elevated incidence. Exposure was represented by the presence and number of Superfund sites in a county. Logistic regression evaluated county-level associations with adjustment for median age, median household income, rurality, and percentage of Black residents. Sixty-eight counties contained at least one Superfund site. A reportable estimate was available for 38 of 68 counties with a site (55.9%) and 25 of 186 counties without a site (13.4%). In the primary adjusted models, site presence (OR = 4.19, 95% CI: 1.66–10.60) and each additional site (OR = 2.76, 95% CI: 1.51–5.04) were associated with higher odds of a reportable estimate. However, after additional adjustment for log-transformed county population, the associations attenuated to the null for site presence (OR = 0.28, 95% CI: 0.02–3.19) and site count (OR = 0.82, 95% CI: 0.31–2.14), and residual spatial autocorrelation was no longer detected. The primary associations therefore appear largely attributable to the concentration of both Superfund sites and reportable cancer counts in populous urban counties. Studies using unsuppressed counts or incidence rates, appropriate population denominators, and spatially resolved exposure measures are needed.

Keywords: early-onset colorectal cancer; county-level ecological study; Superfund sites; Texas; environmental justice

1. Introduction

Despite a general decline in overall colorectal cancer incidence, the incidence of early-onset colorectal cancer (EOCRC), defined as colorectal cancer diagnosed before age 50, has increased for more than 30 years [1]. Between 1990 and 2020, EOCRC incidence increased globally, with important public health implications [2]. Colon and rectal cancer incidence has risen particularly rapidly among adults aged 20–34 years [3]. Approximately 10–15% of EOCRC cases have an identifiable hereditary component, including mismatch repair deficiency [4]; consequently, most cases are considered sporadic. Sporadic EOCRC is often diagnosed in the distal colon or rectum [5] and at an advanced stage, and may exhibit aggressive histopathologic features [6,7]. Although genetic susceptibility and diet are recognized contributors, environmental exposures, including those associated with industrial and hazardous-waste sites, warrant investigation.

Recent epidemiological studies and reviews have focused on environmental exposure and its association with cancers, especially EOCRC [8–10], with several cohort studies utilized for geographic and biological markers of exposure [11,12]. However, many of the cohort studies examined either used mortality rather than cancer incidence as end-points [13] or were not conducted in the USA [14], and most studies enrolled participants with generally higher socioeconomic status, were not enriched with at-risk participants for EO cancers, and did not focus on medically underserved populations in the US [12,13]. Environmental exposures from natural and anthropogenic sources vary within regions of the world, based on proximity to industrial processes, and according to individual-level factors [15,16]. Much of the research has focused on environmental pollution in rural areas. Common environmental exposures include trash burning [16,17], field burning [18], heavy metals [19,20], pesticides and herbicides [21], and occupational exposures [22]. Specifically in southern states [23], industrial and individual processes in rural areas, such as

oil refining, logging, field burning, and diesel exposure, produce environmental exposure that can lead to inflammation, endocrine disruption [24], and increased risk of disease and mortality [25]. Rural populations may have a unique set of environmental exposures compared to their urban counterparts due to the industrial and individual practices in these areas. Combined with lower access to care [26], greater health disparities, and a lack of research into this population in the US [27], there is an urgent need for studies regarding environmental exposures and EOCRC in rural states and communities. On the other hand, environmental injustice in urban areas starkly affects low-income and minority communities, where residents are disproportionately exposed to pollution and contamination. Due to historical segregation and discriminatory housing policies, these communities are often located near sources of air and water pollution, hazardous waste disposal sites, and other environmental health risks [28,29]. The enduring legacy of these policies means that these communities frequently lack the resources to relocate away from these harmful environments. Alarmingly, about 70% of Superfund sites on the National Priorities List (NPL) are situated within one mile of government-assisted housing [30]. This proximity places approximately 77,000 individuals in direct vulnerability to the health hazards emanating from these contaminated sites, underscoring a persistent and systemic disregard for the well-being of these communities.

Environmental contamination from Superfund sites, areas designated by the EPA as critically hazardous, poses significant risks to public health. These sites often emit toxins, including heavy metals and persistent organic pollutants, which permeate air, water, and soil. Evidence suggests that prolonged exposure to pollutants such as lead and arsenic, commonly found at Superfund sites, may disrupt biological processes, leading to inflammation and carcinogenesis. Populations residing near these sites frequently belong to low-income and minority groups, exacerbating health disparities through a convergence of environmental and socioeconomic vulnerabilities. Studies have shown that communities with greater exposure to environmental toxins often experience poorer health outcomes, including higher rates of respiratory conditions, cardiovascular diseases, and certain cancers [31]. These disparities are compounded by limited access to healthcare and inadequate resources to mitigate exposure risks.

The relationship between Superfund site distribution and EOCRC incidence remains underexplored, particularly in Texas [32]. We therefore conducted an exploratory county-level ecological analysis to determine whether the presence and number of Superfund sites were associated with the availability of a non-zero, non-suppressed EOCRC incidence estimate. We hypothesized that counties containing Superfund sites would be more likely to have such an estimate after adjustment for selected demographic and socioeconomic characteristics. Because both exposure and outcome were measured at the county level, this analysis was designed for hypothesis generation and cannot determine whether individual cases lived near, or were exposed to contaminants from, a Superfund site.

2. Materials and Methods

We conducted an ecological analysis of all 254 Texas counties. Superfund sites and their county locations were obtained from the Texas Commission on Environmental Quality (TCEQ) [33]. The exposure variables were (1) the presence of at least one listed site and (2) the number of listed sites in each county. Sixty-eight counties contained at least one site and 186 contained none. The source list included current and former state- and federally administered sites and therefore encompassed sites with differing contaminants, sizes, operational histories, and remediation status. Site coordinates, distance from population centers or county borders, contaminant-specific information, and exposure pathways were not

incorporated. Consequently, county residence and site presence should not be interpreted as individual exposure or residential proximity.

EOCRC data were obtained from the Texas Department of State Health Services Texas Cancer Registry for 2018–2022. County-level incidence estimates based on small numbers were suppressed to protect confidentiality. Because the public-use data did not provide the underlying case counts for suppressed counties, we created a binary data-availability outcome: a county with a non-zero, non-suppressed EOCRC incidence rate estimate was classified as having a reportable estimate; counties with a zero or suppressed estimate were classified as not having a reportable estimate. This measure is influenced by county population size and case counts, is not a clinical or epidemiologic threshold, and is distinct from EOCRC incidence.

County demographic and socioeconomic characteristics were obtained from the U.S. Census Bureau American Community Survey 5-year estimates [34]. Covariates included rurality, median age, median household income, and percentage of Black residents and were selected a priori because they could be associated with both site distribution and outcome reporting. Continuous variables were summarized using means and standard deviations, and categorical variables using counts and column percentages. Group differences were evaluated using t tests and chi-square tests, respectively. Logistic regression estimated odds ratios (ORs) and 95% confidence intervals (CIs) for a reportable EOCRC estimate. Two separate exposure specifications were evaluated: Model A compared counties with at least one Superfund site with counties having no sites, whereas Model B modeled the number of sites continuously, with the OR representing the change in odds per additional site. Each exposure specification was evaluated without adjustment and after adjustment for the selected covariates.

Analyses were conducted using SAS v9.4 (SAS Institute Inc., Cary, NC, USA). Multicollinearity was evaluated using variance inflation factors (VIFs), with values greater than 5 considered potentially important. Model calibration was assessed using the Hosmer–Lemeshow goodness-of-fit test, and discrimination was evaluated using the area under the receiver operating characteristic curve (AUC). Global Moran's I was applied to deviance residuals from each fully adjusted model using a row-standardized first-order queen-contiguity spatial-weights matrix; counties sharing a boundary or vertex were considered neighbors. Statistical significance was evaluated using 999 permutations.

Because the binary outcome was determined partly by whether a county accumulated enough EOCRC cases for its estimate to be reported, we conducted sensitivity analyses that additionally adjusted for the natural logarithm of total county population. Multicollinearity, calibration, discrimination, and residual spatial autocorrelation were reassessed after adding county population. If no significant residual spatial autocorrelation remained, an additional spatial regression model was not fitted because the observed spatial dependence had been accounted for by the measured population covariate.

3. Results

The analysis included all 254 Texas counties: 68 with at least one Superfund site and 186 without a site. Table 1 compares county-level demographic and socioeconomic characteristics by site presence. A non-zero, non-suppressed EOCRC estimate was available for 38 of 68 counties with a site (55.9%) and 25 of 186 counties without a site (13.4%; $p < 0.001$). Counties with sites had a lower mean median age (37.2 vs. 40.3 years; $p < 0.001$) and a higher mean percentage of Black residents (9.8% vs. 4.8%; $p < 0.001$). Median household income was $79,393 in counties with sites and $74,538 in counties without sites ($p = 0.05$). The percentages impoverished and uninsured did not differ significantly between

groups. Rural counties comprised 44.1% of counties with sites and 74.2% of counties without sites ($p < 0.001$).

Table 1. County characteristics according to the presence of at least one Superfund site.

County Characteristics	Superfund Site		*p*-Value
	Yes (N = 68)	No (N = 186)	
Dichotomous Variables	N (%)	N (%)	
Non-zero, non-suppressed EOCRC estimate	38 (55.9)	25 (13.4)	<0.001
Rural	30 (44.1)	138 (74.2)	<0.001
Continuous Variables	Mean (Std)	Mean (Std)	
Median Household Income	79,393.2 (17,311.1)	74,537.8 (17,591.9)	0.05
% Population Impoverished	11.4 (4.6)	11.5 (4.6)	0.88
% Black	9.8 (7.7)	4.8 (5.2)	<0.001
% Employed as Civilian Labor	56.4 (6.4)	53.8 (7.6)	0.02
% Received Food Stamps	12.8 (5.2)	12.1 (6.8)	0.35
% >25 years With High School Diploma	16.0 (5.9)	17.4 (8.4)	0.13
% Hispanics	34.2 (21.7)	36.6 (23.9)	0.46
Median age	37.2 (4.8)	40.3 (6.5)	<0.001
Median age of Females	38.1 (5.0)	41.4 (7.2)	<0.001
Median age of Males	36.3 (4.8)	39.4 (6.6)	<0.001
% Population Without Health Insurance	17.5 (4.1)	17.4 (5.4)	0.90
% Females	49.3 (2.3)	48.6 (3.8)	0.07
% >65 years old	15.9 (4.6)	19.4 (4.6)	<0.001
% Unemployed	2.9 (1.1)	2.6 (1.4)	0.07

Values are n (column %) for categorical variables and mean (standard deviation) for continuous variables. *p* values are based on t tests for continuous variables and chi-square tests for categorical variables. The EOCRC outcome denotes a non-zero, non-suppressed county estimate and is not an incidence threshold.

Table 2 presents two distinct logistic regression specifications. Model A treats Superfund exposure as binary and compares counties with at least one site with counties having no sites. Model B treats the number of sites as a continuous exposure and reports the OR per additional site. In Model A, counties with at least one site had higher odds of a nonzero, non-suppressed EOCRC estimate in the unadjusted analysis (OR = 8.16, 95% CI: 4.31–15.43) and after adjustment for median age, median household income, rurality, and percentage of Black residents (adjusted OR = 4.19, 95% CI: 1.66–10.60). In Model B, each additional site was associated with higher odds in the unadjusted (OR = 3.32, 95% CI: 2.21–4.98) and adjusted models (adjusted OR = 2.76, 95% CI: 1.51–5.04). These estimates concern counties and the reportability of county estimates; they do not estimate individual EOCRC risk.

No evidence of problematic multicollinearity was observed. VIF values ranged from 1.10 to 1.37 in Model A and from 1.08 to 1.39 in Model B, well below the prespecified threshold of 5. The Hosmer–Lemeshow test did not indicate lack of calibration for Model A (chi-square = 11.59, 8 df, $p = 0.170$) or Model B (chi-square = 8.75, 8 df, $p = 0.364$). The models demonstrated strong discrimination, with AUC values of 0.931 and 0.937, respectively.

Global Moran's I identified positive residual spatial autocorrelation in the primary adjusted Model A (I = 0.111, permutation $p = 0.002$) and Model B (I = 0.104, permutation $p = 0.005$), indicating that geographic dependence among neighboring counties was not fully accounted for by the original covariates.

Table 2. Primary logistic regression models for a non-zero, non-suppressed county EOCRC incidence estimate using two Superfund exposure specifications.

Model Term	Model A: Site Presence (≥1 Site vs. 0 Sites)	Model B: Site Count (OR per Additional Site)
	OR (95% CI)	OR (95% CI)
Unadjusted exposure model		
Superfund exposure	8.16 (4.31, 15.43)	3.32 (2.21, 4.98)
Adjusted model		
Superfund exposure	4.19 (1.66, 10.60)	2.76 (1.51, 5.04)
Median age (yr)	0.81 (0.72, 0.90)	0.81 (0.72, 0.90)
Median household income	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
Rural vs. urban	0.07 (0.03, 0.18)	0.08 (0.03, 0.20)
% Black vs. other	1.09 (1.01, 1.17)	1.07 (0.99, 1.16)

OR, odds ratio; CI, confidence interval. Model A defines exposure as the presence of one or more Superfund sites and compares exposed counties with counties having no sites. Model B defines exposure as the number of Superfund sites in the county and reports the OR associated with one additional site. The two columns therefore represent separate regression models and not sequential levels of the same exposure. Adjusted models include median age, median household income, rurality, and percentage of Black residents.

Population-adjusted sensitivity analyses materially changed the findings (Table 3). After additional adjustment for log-transformed total county population, neither Superfund site presence (adjusted OR = 0.28, 95% CI: 0.02–3.19; $p = 0.306$) nor site count (adjusted OR per additional site = 0.82, 95% CI: 0.31–2.14; $p = 0.680$) remained associated with the outcome. Rurality was also no longer significant in Model A (OR = 2.55, 95% CI: 0.27–24.39; $p = 0.417$) or Model B (OR = 2.31, 95% CI: 0.25–20.98; $p = 0.458$). No significant residual spatial autocorrelation remained in Model A (Moran's I = −0.020, permutation $p = 0.642$) or Model B (Moran's I = −0.011, permutation $p = 0.829$). These findings indicate that county population accounted for much of the association and spatial pattern observed in the primary models.

Table 3. Population-adjusted sensitivity analyses and residual spatial autocorrelation.

Exposure Specification	Adjusted OR (95% CI)	*p*-Value	Moran's I	Permutation *p*-Value
Model A: ≥1 site vs. 0 sites	0.28 (0.02–3.19)	0.306	−0.020	0.642
Model B: Per additional site	0.82 (0.31–2.14)	0.680	−0.011	0.829

OR, odds ratio; CI, confidence interval. Both sensitivity models additionally adjust for the natural logarithm of total county population, median age, median household income, rurality, and percentage of Black residents. Moran's I was calculated from the deviance residuals of each population-adjusted model using 999 permutations.

4. Discussion

In this exploratory ecological analysis, Texas counties containing Superfund sites were more likely than counties without sites to have a non-zero, non-suppressed EOCRC estimate. The associations remained positive after adjustment for selected county characteristics but attenuated to the null after additional adjustment for total county population. This attenuation should not be interpreted as evidence that less-populated counties have lower EOCRC incidence or that proximity to Superfund sites is unrelated to EOCRC risk. County population strongly influences whether enough EOCRC cases accumulate to produce a reportable incidence estimate. A sparsely populated county may have an elevated EOCRC incidence rate but still have its estimate suppressed because of a small number of cases. At the same time, populous urban counties are more likely to contain Superfund sites because of the historical concentration of population, industrial activities, and waste-generating facilities in metropolitan areas. Consequently, the observed association between Superfund-

site measures and EOCRC estimate reportability cannot be separated from population size and urban concentration using the present outcome measure. The reported odds ratios therefore should not be interpreted as estimates of individual exposure effects or causal risk.

These findings nevertheless have potentially important environmental-justice implications. The concentration of Superfund sites in populous urban counties means that large numbers of residents, and potentially socioeconomically marginalized or historically overburdened communities, may live in areas affected by hazardous-waste sites. Conversely, suppression of cancer estimates in sparsely populated counties may obscure elevated disease burdens in rural communities and create a form of data invisibility. Thus, the absence of a reportable county estimate should not be interpreted as the absence of EOCRC or environmental risk. Although counties with Superfund sites had a higher mean percentage of Black residents, these county-level patterns do not establish that Black residents or individuals with EOCRC experienced greater site-related exposure. Future studies should use unsuppressed case counts or incidence data with appropriate population denominators, finer geographic units, residential-distance measures, and individual-level exposure information to determine whether communities near Superfund sites experience elevated EOCRC incidence. A strength of this study is its statewide inclusion of all 254 counties and its integration of publicly available environmental, demographic, and cancer-registry information. Its limitations remain substantial. First, the outcome is not EOCRC incidence; it identifies whether a county accumulated sufficient cases to produce a reportable estimate and is therefore strongly determined by the underlying population at risk. Although population adjustment characterized this limitation, it cannot transform the binary measure into an incidence outcome. Second, the ecological design precludes individual-level inference: the analysis cannot establish that any person with EOCRC lived near a site or encountered a site-related contaminant. Third, county-level site presence is a crude exposure surrogate that does not capture distance, cross-border exposure, site area, duration, contaminant type or concentration, remediation status, migration, or exposure pathways. Fourth, obesity, diabetes, smoking, alcohol use, diet, physical activity, healthcare utilization, screening, and other urban environmental exposures were unavailable, leaving potential residual confounding. Finally, although significant spatial autocorrelation was present in the primary models, it was no longer detected after adjustment for county population. We therefore did not fit an additional spatial regression model; future studies using actual counts or incidence rates should nevertheless employ spatial count or rate models.

Most counties with Superfund sites were urban, consistent with the concentration of industrial activity and larger populations in metropolitan areas. The disappearance of the Superfund and rurality associations after population adjustment supports an important alternative explanation: in the primary models, Superfund site presence functioned mainly as a marker of urban population concentration, reporting capacity, healthcare access, and case detection rather than as evidence of an independent environmental contribution to EOCRC. Counties with sites also had a higher mean percentage of Black residents, raising environmental justice questions; however, these county-level co-occurrences do not establish that Black residents or EOCRC cases experienced greater site-related exposure.

Prior studies have used more spatially resolved exposure definitions than were possible with the public county-level data used here. For example, a Massachusetts study evaluated residence within 3 km of a Department of Defense Superfund site [35]; spatial scan statistics have been used to compare cancer patterns with densities of EPA-designated sites [36]; and studies in Florida, Montana, and Kentucky have considered hazard levels, contaminated water, or residential proximity [37–39]. These approaches underscore an important limitation of our county-presence measure. Applying a residential buffer to county-level cancer summaries would create false precision because case locations were

unavailable, and a site-centered buffer crossing county lines could not be linked validly to individual cases. Access to geocoded case residences and residential histories is therefore necessary for defensible distance-based analysis.

The present analysis does not identify the contaminants responsible for the county-level association. The TCEQ source list includes sites with heterogeneous contaminant profiles and remediation histories, and complete standardized contaminant data were not available for all sites. Volatile organic compounds, metals, polychlorinated biphenyls, pesticides, and herbicides may occur at some sites, but their presence and pathways cannot be assumed for every site. Future work should construct site-specific exposure metrics incorporating contaminant identity, environmental media, toxicity, remediation status, duration, dispersion, and population proximity.

This study is best viewed as an evaluation of a county-level reporting pattern rather than an analysis of EOCRC incidence. A stronger study would use unsuppressed age-specific case counts or incidence rates with appropriate population denominators; link geocoded residential histories to site coordinates; model distance, transport pathways, and exposure duration; incorporate site status and contaminants; adjust for behavioral and clinical risk factors; and use spatial count or rate models. These improvements would avoid the information loss and population dependence introduced by the present binary outcome.

5. Conclusions

In this exploratory ecological study, Superfund site presence and site count were associated with the availability of a non-zero, non-suppressed county-level EOCRC estimate in the primary models, but these associations attenuated after adjustment for total county population. This attenuation underscores that the outcome reflects both the occurrence of EOCRC and the population-dependent ability to generate a reportable estimate. It does not demonstrate the absence of EOCRC risk in rural counties or the absence of health concerns in urban communities near Superfund sites. The geographic concentration of hazardous-waste sites in populous communities and the potential underrepresentation of disease burden in counties with suppressed estimates warrant further investigation using unsuppressed case counts, population denominators, finer spatial resolution, and individual-level exposure information.

Author Contributions: Research concept initiation: L.J.S., E.H.H., S.G.Y., P.-C.H., M.L., H.K., M.B., K.S. and S.-M.H.; Statistical Analysis and Result Summary: L.J.S., J.E. and M.L.; Initial Draft: L.J.S. and J.E.; Review and Edit: L.J.S., E.H.H., S.G.Y., P.-C.H., M.L. and S.-M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the National Cancer Institute, Cancer Center Support Grant P30CA142543.

Institutional Review Board Statement: Ethical review and approval were waived for this study because only publicly available, aggregated, population-based summary data were used to conduct the ecologic analyses. No personal identifiers were used in the analyses.

Informed Consent Statement: Not applicable. This study used only publicly available, aggregated, population-based summary data to conduct the ecologic analyses. No personal identifiers were used in the analyses.

Data Availability Statement: This study utilized publicly available data. All data sources are provided in the reference citation.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EOCRC	Early-onset colorectal cancer
ANOVA	Analysis of Variance
OR	Odds Ratio
CI	Confidence Interval
NPL	National Priorities List
EPA	Environmental Protection Agency
DSHS	Texas Department of State Health Services
SEER	Surveillance, Epidemiology and End Results
ACS	American Community Survey
TCEQ	Texas Commission on Environmental Quality
PCB	polychlorinated biphenyl
VIF	Variance inflation factor
AUC	Area under the receiver operating characteristic curve

References

1. Siegel, R.L.; Wagle, N.S.; Cercek, A.; Smith, R.A.; Jemal, A. Colorectal cancer statistics, 2023. *CA Cancer J. Clin.* **2023**, *73*, 233–254. [CrossRef]
2. Burwell-Naney, K.; Zhang, H.; Samantapudi, A.; Jiang, C.; Dalemarre, L.; Rice, L.; Williams, E.; Wilson, S. Spatial disparity in the distribution of superfund sites in South Carolina: An ecological study. *Environ. Health* **2013**, *12*, 96. [CrossRef]
3. Eng, C.; Hochster, H. Early-Onset Colorectal Cancer: The Mystery Remains. *J. Natl. Cancer Inst.* **2021**, *113*, 1608–1610. [CrossRef]
4. Salovaara, R.; Loukola, A.; Kristo, P.; Kääriäinen, H.; Ahtola, H.; Eskelinen, M.; Härkönen, N.; Julkunen, R.; Kangas, E.; Ojala, S.; et al. Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J. Clin. Oncol.* **2000**, *18*, 2193–2200. [CrossRef]
5. You, Y.N.; Xing, Y.; Feig, B.W.; Chang, G.J.; Cormier, J.N. Young-onset colorectal cancer: Is it time to pay attention? *Arch. Intern. Med.* **2012**, *172*, 287–289. [CrossRef]
6. Ballester, V.; Rashtak, S.; Boardman, L. Clinical and molecular features of young-onset colorectal cancer. *World J. Gastroenterol.* **2016**, *22*, 1736–1744. [CrossRef]
7. Ahnen, D.J.; Wade, S.W.; Jones, W.F.; Sifri, R.; Silveiras, J.M.; Greenamyer, J.; Guiffre, S.; Axilbund, J.; Spiegel, A.; You, Y.N. The increasing incidence of young-onset colorectal cancer: A call to action. *Mayo Clin. Proc.* **2014**, *89*, 216–224. [CrossRef]
8. AlZaabi, A.; Younus, H.A.; Al-Reasi, H.A.; Al-Hajri, R. Could environmental exposure and climate change Be a key factor in the rising incidence of early onset colorectal cancer? *Heliyon* **2024**, *10*, e35935. [CrossRef]
9. Chen, A.Y.; Matich, E.K.; Laryea, J.; Hsu, P.C.; Su, L.J. A Case-Control Study of Dietary Choline Intake and Risk of Colorectal Cancer Modified by Dietary B-Vitamin Intake. *Nutrients* **2024**, *16*, 4200. [CrossRef]
10. Hofseth, L.J.; Hebert, J.R.; Chanda, A.; Chen, H.; Love, B.L.; Pena, M.M.; Murphy, E.A.; Sajish, M.; Sheth, A.; Buckhaults, P.J.; et al. Early-onset colorectal cancer: Initial clues and current views. *Nat. Rev. Gastroenterol. Hepatol.* **2020**, *17*, 352–364. [CrossRef]
11. Terry, M.B.; Michels, K.B.; Brody, J.G.; Byrne, C.; Chen, S.; Jerry, D.J.; Malecki, K.M.C.; Martin, M.B.; Miller, R.L.; Neuhausen, S.L.; et al. Environmental exposures during windows of susceptibility for breast cancer: A framework for prevention research. *Breast Cancer Res.* **2019**, *21*, 96. [CrossRef]
12. Zeinomar, N.; Oskar, S.; Kehm, R.D.; Sahebzeda, S.; Terry, M.B. Environmental exposures and breast cancer risk in the context of underlying susceptibility: A systematic review of the epidemiological literature. *Environ. Res.* **2020**, *187*, 109346. [CrossRef]
13. Turner, M.C.; Krewski, D.; Diver, W.R.; Pope, C.A.; Burnett, R.T.; Jerrett, M.; Marshall, J.D.; Gapstur, S.M. Ambient Air Pollution and Cancer Mortality in the Cancer Prevention Study II. *Environ. Health Perspect.* **2017**, *125*, 087013. [CrossRef]
14. Jiang, F.; Zhao, J.; Sun, J.; Chen, W.; Zhao, Y.; Zhou, S.; Yuan, S.; Timofeeva, M.; Law, P.J.; Larsson, S.C.; et al. Impact of ambient air pollution on colorectal cancer risk and survival: Insights from a prospective cohort and epigenetic Mendelian randomization study. *EBioMedicine* **2024**, *103*, 105126. [CrossRef]
15. International Agency for Research on Cancer. Outdoor Air Pollution. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*; International Agency for Research on Cancer: Lyon, France, 2015.
16. Vreeland, H.; Schauer, J.J.; Russell, A.G.; Marshall, J.D.; Fushimi, A.; Jain, G.; Sethuraman, K.; Verma, V.; Tripathi, S.N.; Bergin, M.H. Chemical characterization and toxicity of particulate matter emissions from roadside trash combustion in urban India. *Atmos. Environ.* **2016**, *147*, 22–30. [CrossRef]
17. Loomis, D.; Grosse, Y.; Lauby-Secretan, B.; El Ghissassi, F.; Bouvard, V.; Benbrahim-Tallaa, L.; Guha, N.; Baan, R.; Mattock, H.; Straif, K. The carcinogenicity of outdoor air pollution. *Lancet Oncol.* **2013**, *14*, 1262–1263. [CrossRef]
18. Estrellan, C.R.; Iino, F. Toxic emissions from open burning. *Chemosphere* **2010**, *80*, 193–207. [CrossRef]

19. Kresovich, J.K.; Erdal, S.; Chen, H.Y.; Gann, P.H.; Argos, M.; Rauscher, G.H. Metallic air pollutants and breast cancer heterogeneity. *Environ. Res.* **2019**, *177*, 108639. [CrossRef]
20. White, A.J.; O'Brien, K.M.; Niehoff, N.M.; Carroll, R.; Sandler, D.P. Metallic Air Pollutants and Breast Cancer Risk in a Nationwide Cohort Study. *Epidemiology* **2019**, *30*, 20–28. [CrossRef]
21. Matich, E.K.; Laryea, J.A.; Seely, K.A.; Stahr, S.; Su, L.J.; Hsu, P.C. Association between pesticide exposure and colorectal cancer risk and incidence: A systematic review. *Ecotoxicol. Environ. Saf.* **2021**, *219*, 112327. [CrossRef]
22. O'sUllivan, D.E.; Sutherland, R.L.; Town, S.; Chow, K.; Fan, J.; Forbes, N.; Heitman, S.J.; Hilsden, R.J.; Brenner, D.R. Risk Factors for Early-Onset Colorectal Cancer: A Systematic Review and Meta-analysis. *Clin. Gastroenterol. Hepatol.* **2022**, *20*, 1229–1240 e5. [CrossRef]
23. Johnston, J.E.; Chau, K.; Franklin, M.; Cushing, L. Environmental Justice Dimensions of Oil and Gas Flaring in South Texas: Disproportionate Exposure among Hispanic communities. *Environ. Sci. Technol.* **2020**, *54*, 6289–6298. [CrossRef]
24. Darbre, P.D. Overview of air pollution and endocrine disorders. *Int. J. Gen. Med.* **2018**, *11*, 191–207. [CrossRef]
25. Turner, M.C.; Andersen, Z.J.; Baccarelli, A.; Diver, W.R.; Gapstur, S.M.; Pope, C.A.; Prada, D.; Samet, J.; Thurston, G.; Cohen, A. Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations. *CA Cancer J. Clin.* **2020**, *70*, 460–479. [CrossRef]
26. Borders, T.F.; Morgan, Z.J.; Peterson, L.E. Colorectal Cancer Screening in Rural and Urban Primary Care Practices Amid Implementation of the Medicare Access and CHIP Reauthorization Act. *J. Prim. Care Community Health* **2023**, *14*, 21501319231177552. [CrossRef]
27. Muthukrishnan, M.; Sutcliffe, S.; Hunleth, J.M.; Wang, J.S.; Colditz, G.A.; James, A.S. Conducting a randomized trial in rural and urban safety-net health centers: Added value of community-based participatory research. *Contemp. Clin. Trials Commun.* **2018**, *10*, 29–35. [CrossRef]
28. Estien, C.O.; Wilkinson, C.E.; Morello-Frosch, R.; Schell, C.J. Historical Redlining Is Associated with Disparities in Environmental Quality across California. *Environ. Sci. Technol. Lett.* **2024**, *11*, 54–59. [CrossRef]
29. Woo, B.; Kravitz-Wirtz, N.; Sass, V.; Crowder, K.; Teixeira, S.; Takeuchi, D.T. Residential Segregation and Racial/Ethnic Disparities in Ambient Air Pollution. *Race Soc. Probl.* **2019**, *11*, 60–67. [CrossRef]
30. U.S. EPA. *Population Surrounding 1881 Superfund Sites. Data Collected Includes: (1) Superfund Site Information from SEMS as of the End of FY2022 and Site Boundary Data from EPA Shared Enterprise Geodata Services (SEGS) 2023 and (2) Population Data from the 2017-2021 American Community Survey*; U.S. EPA, Office of Land and Emergency Management: Washington, DC, USA, 2023. Available online: https://www.epa.gov/system/files/documents/2023-08/FY22%20Population%20Estimates%20Superfund%20Final.pdf (accessed on 15 August 2025).
31. Kiaghadi, A.; Rifai, H.S.; Dawson, C.N. The presence of Superfund sites as a determinant of life expectancy in the United States. *Nat. Commun.* **2021**, *12*, 1947. [CrossRef]
32. Alves, B. Number of Hazardous Waste Sites in the U.S. 2025, by State. Statistica Inc. Available online: https://www.statista.com/statistics/1147665/number-of-hazardous-waste-sites-in-the-united-states/#statisticContainer (accessed on 15 August 2025).
33. Texas Commission on Environmental Quality. *Superfund Sites in Texas by Site Name*; Texas Commission on Environmental Quality: Austin, TX, USA, 2024.
34. U.S. Census Bureau. *American Community Survey Data*; U.S. Census Bureau: Washington, DC, USA, 2023. Available online: https://www.census.gov/programs-surveys/acs/data.html (accessed on 15 August 2025).
35. Ozonoff, D.; Aschengrau, A.; Coogan, P. Cancer in the vicinity of a Department of Defense superfund site in Massachusetts. *Toxicol. Ind. Health* **1994**, *10*, 119–141. [CrossRef]
36. Amin, R.; Nelson, A.; McDougall, S. A Spatial Study of the Location of Superfund Sites and Associated Cancer Risk. *Stat. Public Policy* **2018**, *5*, 1–9. [CrossRef]
37. Davis, B.; McDermott, S.; McCarter, M.; Ortaglia, A. Population-based mortality data suggests remediation is modestly effective in two Montana Superfund counties. *Environ. Geochem. Health* **2019**, *41*, 803–816. [CrossRef]
38. Kirpich, A.; Leary, E. Superfund Locations and Potential Associations with Cancer Incidence in Florida. *Stat. Public Policy* **2017**, *4*, 1–9. [CrossRef]
39. Stone, R.; Webber, W.B. Male-female differences in the prevalence of non-hodgkin lymphoma and residential proximity to superfund sites in kentucky. *J. Environ. Health* **2018**, *81*, 16–24.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.