

Rural vs. Metropolitan Health Disparities Across Texas Counties

A prespecified county-level comparison of heart disease deaths, lung cancer incidence and six CDC PLACES estimates between metropolitan and nonmetropolitan counties.

Author

Agamvir Singh
Independent Researcher · Vellora

Version

1.2
Updated 2026-10-24

Published

2026-10-24

Type

Ecological, cross-sectional comparison of counties; not peer reviewed

Data years

Deaths 2022–2024; cancer diagnoses 2019–2023;
BRFSS 2023 (PLACES 2025); RUCC 2023

Online report, with the interactive figure, data and code: vellora-health.pages.dev/research/reports/rural-metropolitan-texas/

Abstract

Background. National surveillance has reported higher age-adjusted death rates and higher lung cancer incidence in nonmetropolitan than in metropolitan U.S. counties. ^[3, 4] This study compares Texas’s 86 metropolitan and 168 nonmetropolitan counties on eight verified county measures.

Methods. Ecological, cross-sectional comparison using USDA Rural-Urban Continuum Codes 2023 (codes 1–3 metropolitan, 4–9 nonmetropolitan). A plan written before any comparison was computed defined a primary family of two observed rates, heart disease deaths among adults 35 and over (2022–2024) and lung cancer incidence

(2019–2023), and a secondary family of six CDC PLACES age-adjusted estimates (BRFSS 2023). For each, we report group medians, the Hodges–Lehmann shift (nonmetropolitan minus metropolitan) with a 95% confidence interval, the probability of superiority and a Mann–Whitney test, with Benjamini–Hochberg adjustment within each family. Suppressed values were excluded, never imputed.

Results. Heart disease death rates were higher in nonmetropolitan counties (median 415.1 per 100,000 against 361.2 per 100,000; shift +45.4 per 100,000 (95% CI +25.7 to +65.0); adjusted $p < 0.0001$). Lung cancer incidence showed no clear difference (median 52.9 against 50.8; shift +3.5 per 100,000 (95% CI –0.9 to +7.9); adjusted $p = 0.1340$), but its rate was suppressed for 43 of 168 nonmetropolitan and 8 of 86 metropolitan counties. All six PLACES estimates were higher in nonmetropolitan counties (probability of superiority 0.61–0.73; adjusted $p \leq 0.0033$).

Conclusions. Nonmetropolitan Texas counties tended to have higher heart disease death rates and higher modelled prevalence of cardiometabolic conditions and risk behaviours. The lung cancer comparison is inconclusive and limited by suppression in small counties. These are differences between counties, not evidence that where people live causes them, and they do not describe individuals.

1. Introduction

CDC has reported higher age-adjusted death rates from leading causes, including heart disease, in nonmetropolitan than in metropolitan areas of the United States,^[3] and higher lung cancer incidence in nonmetropolitan counties.^[4] Texas spans 254 counties, from counties in metropolitan areas of more than a million people to counties with a few dozen residents, so it is a useful place to ask whether the national pattern appears in current county data.

This report uses USDA’s county-level metropolitan and nonmetropolitan categories. They are not the same as “urban” and “rural”: USDA notes that most counties, metropolitan or not, contain both urban and rural populations, and that metropolitan areas include substantial rural territory.^[2] The word “rural” in the title follows common usage; the analysis itself compares metropolitan and nonmetropolitan counties.

Vellora’s earlier report [Geographic Variation in Cardiovascular Health Across Texas Counties](#) compared the two groups on eight cardiovascular measures. This study reuses

some of the same verified data and reproduces those comparisons exactly where they overlap, and adds an observed cancer outcome, a prespecified primary family, effect sizes in each measure’s own units with confidence intervals, an analysis of missing data and sensitivity analyses.

Research question: How do selected health outcomes and risk factors differ between metropolitan and nonmetropolitan Texas counties?

2. Data and methods

2.1 Groups

Counties were classified with the USDA Economic Research Service Rural-Urban Continuum Codes 2023 (RUCC),^[1] matched on five-digit FIPS codes: codes 1–3 metropolitan and 4–9 nonmetropolitan. The stored codes were checked against a fresh download of the USDA file on 2026-10-23: all 254 Texas rows matched, giving 86 metropolitan and 168 nonmetropolitan counties (Table 1).

Table 1. Texas counties by Rural-Urban Continuum Code, 2023			
Code	Description	Group	Counties
1	Metropolitan: in a metro area of 1 million population or more	Metropolitan	34
2	Metropolitan: in a metro area of 250,000 to 1 million population	Metropolitan	35
3	Metropolitan: in a metro area of fewer than 250,000 population	Metropolitan	17
4	Nonmetropolitan: urban population of 20,000 or more, adjacent to a metro area	Nonmetropolitan	10
5	Nonmetropolitan: urban population of 20,000 or more, not adjacent to a metro area	Nonmetropolitan	6
6	Nonmetropolitan: urban population of 5,000 to 20,000, adjacent to a metro area	Nonmetropolitan	42
7	Nonmetropolitan: urban population of 5,000 to 20,000, not adjacent to a metro area	Nonmetropolitan	18

Code	Description	Group	Counties
8	Nonmetropolitan: urban population of fewer than 5,000, adjacent to a metro area	Nonmetropolitan	46
9	Nonmetropolitan: urban population of fewer than 5,000, not adjacent to a metro area	Nonmetropolitan	46

2.2 Indicators

Eight indicators were selected for relevance, county coverage and data quality (Table 2). The primary family contains the two observed rates, counted from death certificates and the cancer registry. The six PLACES measures form a secondary family because they are small-area model estimates whose inputs include county demographics, so part of any difference between the groups may come from the model itself. ^[9] Stroke prevalence and other PLACES and cancer measures were not selected, to keep the comparison focused. Crude and age-adjusted values are never combined.

Table 2. Indicators, sources and adjustment			
Indicator	Family	Source and period	Unit and adjustment
Heart disease death rate, adults 35+	Primary	CDC heart disease mortality (DHDSP), Deaths 2022–2024 ^[5, 6]	Deaths per 100,000 adults 35+; Age-standardized (2000 U.S. standard); spatially smoothed
Lung and bronchus cancer incidence	Primary	State Cancer Profiles (Texas Cancer Registry), Diagnoses 2019–2023 ^[7]	New cases per 100,000 people; Age-adjusted (2000 U.S. standard)
Coronary heart disease among adults	Secondary	CDC PLACES 2025, BRFSS 2023 ^[8, 9]	% of adults; Age-adjusted prevalence, adults 18+ (model-based)
High blood pressure among adults	Secondary	CDC PLACES 2025, BRFSS 2023 ^[8, 9]	% of adults; Age-adjusted prevalence, adults 18+ (model-based)
Diagnosed diabetes among adults	Secondary	CDC PLACES 2025, BRFSS 2023 ^[8, 9]	% of adults; Age-adjusted prevalence, adults 18+ (model-based)
Obesity among adults	Secondary	CDC PLACES 2025, BRFSS 2023 ^[8, 9]	% of adults; Age-adjusted prevalence, adults 18+ (model-based)

Indicator	Family	Source and period	Unit and adjustment
			based)
Current cigarette smoking among adults	Secondary	CDC PLACES 2025, BRFSS 2023 ^[8, 9]	% of adults; Age-adjusted prevalence, adults 18+ (model-based)
No leisure-time physical activity among adults	Secondary	CDC PLACES 2025, BRFSS 2023 ^[8, 9]	% of adults; Age-adjusted prevalence, adults 18+ (model-based)

Age adjustment. The death rate and the cancer rate are adjusted to the 2000 U.S. standard population (adults 35 and over for deaths, all ages for cancer). ^[5, 7] PLACES age-adjusted estimates cover adults 18 and over; CDC’s PLACES documentation does not name their standard population. ^[9, 10] Each indicator is compared only with itself across the two groups, never with another indicator.

2.3 Statistical analysis

The analysis followed a plan committed to the project’s version history before any group difference was computed (see Reproducibility). Differences are always nonmetropolitan minus metropolitan. For each indicator and group we report the number of counties with a published value, the number suppressed, the median, interquartile range (IQR), mean, standard deviation and range.

- **Difference in medians:** the nonmetropolitan median minus the metropolitan median.
- **Hodges–Lehmann shift:** the median of every difference between a nonmetropolitan and a metropolitan county, a robust estimate of how far one distribution sits from the other, with a distribution-free 95% confidence interval. ^[12, 13]
- **Probability of superiority:** the chance that a randomly chosen nonmetropolitan county has a higher value than a randomly chosen metropolitan county (0.5 means no tendency either way), and the equivalent rank-biserial correlation ($2 \times \text{probability} - 1$).
- **Mann–Whitney U test**, two-sided, with tie and continuity corrections. ^[11] p-values are adjusted with the Benjamini–Hochberg procedure separately within the primary family (2 tests) and the secondary family (6 tests). ^[14] Both raw and adjusted values are reported.

A small p-value says a difference is unlikely to be due to chance alone, not that it is large or important; a non-significant result does not show that the groups are the same. Counties

are unweighted. A population-weighted analysis was not performed: the values are age-adjusted rates and model-based prevalence estimates, and weighting them by 2020 population would not produce a valid pooled rate without age-specific counts, which are not in the verified data.

2.4 Missing data and exploratory analyses

Suppressed values were excluded and listed with the source's reason; nothing was imputed or replaced with zero. For each indicator with suppressed values, Fisher's exact test compares how often values are suppressed in each group. Three exploratory analyses (unadjusted p-values, not used for the conclusions) were planned: E1 repeats the comparison with crude rather than age-adjusted PLACES prevalence; E2 correlates each indicator with the RUCC code from 1 to 9 (Spearman); E3 compares nonmetropolitan counties adjacent to a metropolitan area (codes 4, 6, 8) with nonadjacent ones (5, 7, 9).

3. Results

3.1 Counties and sample sizes

The 86 metropolitan counties held 90% of Texans in 2020 (26,249,929 people; median county 92,260.5), and the 168 nonmetropolitan counties 10% (2,895,576; median county 11,480.5). Every indicator except lung cancer incidence was published for all counties but Loving County (nonmetropolitan). Table 3 gives each indicator's own sample sizes.

Table 3. Counties with a published value and county medians, by group

Indicator	Metropolitan n (suppressed)	Metropolitan median (IQR)	Nonmetropolitan n (suppressed)	Nonmetropolitan median (IQR)
Heart disease deaths (per 100,000)	86 (0)	361.2 (316.6– 414.1)	167 (1)	415.1 (366.0–457.6)
Lung cancer incidence (per 100,000)	78 (8)	50.8 (42.0– 60.9)	125 (43)	52.9 (44.1–67.3)
Coronary heart disease (%)	86 (0)	5.6 (5.2–5.9)	167 (1)	6.1 (5.8–6.5)
High blood pressure (%)	86 (0)	33.0 (31.7–34.1)	167 (1)	33.9 (32.6–35.3)
Diabetes (%)	86 (0)	11.6 (10.8–13.0)	167 (1)	12.5 (11.3–13.9)
Obesity (%)	86 (0)	36.4 (34.6– 39.0)	167 (1)	38.0 (36.5–39.8)
Current smoking (%)	86 (0)	14.1 (12.6–14.9)	167 (1)	15.9 (14.6–17.0)
No leisure-time activity (%)	86 (0)	27.8 (25.4– 30.8)	167 (1)	30.5 (28.1–33.8)

3.2 Primary family: observed rates

Heart disease deaths. The median nonmetropolitan county's death rate was 415.1 per 100,000, against 361.2 per 100,000 for metropolitan counties. The Hodges–Lehmann shift was +45.4 per 100,000 (95% CI +25.7 to +65.0), and a randomly chosen nonmetropolitan county had the higher rate with probability 0.67 (Mann–Whitney $p < 0.0001$; adjusted $p < 0.0001$).

Lung cancer incidence. Among counties with a published rate, the medians were 52.9 (nonmetropolitan, $n = 125$) and 50.8 (metropolitan, $n = 78$) per 100,000. The shift was +3.5 per 100,000 (95% CI -0.9 to $+7.9$); the interval includes zero and the difference was not statistically significant (adjusted $p = 0.1340$). This does not show that the groups have the same incidence: see 3.4 on suppression.

Figure 1. County values by metropolitan status, all eight indicators. Each dot is one county with a published value; the dark line is the group median and the shaded band its interquartile range. Suppressed counties are not shown (Table 5).

Heart disease death rate, adults 35+

Metropolitan · n = 86, median 361.2



Nonmetropolitan · n = 167, median 415.1



200 400 600 800 1000

Heart disease deaths (per 100,000, Deaths 2022–2024)

Nonmetropolitan minus metropolitan: median difference +53.9; Hodges–Lehmann shift +45.4 (95% CI +25.7 to +65.0) per 100,000; probability of superiority 0.67; BH-adjusted $p < 0.0001$ (primary family).

Lung and bronchus cancer incidence

Metropolitan · n = 78, median 50.8



Nonmetropolitan · n = 125, median 52.9



20 40 60 80 100

Lung cancer incidence (per 100,000, Diagnoses 2019–2023)

Nonmetropolitan minus metropolitan: median difference +2.1; Hodges–Lehmann shift +3.5 (95% CI –0.9 to +7.9) per 100,000; probability of superiority 0.56; BH-adjusted $p = 0.1340$ (primary family).

Coronary heart disease among adults

Metropolitan · n = 86, median 5.6



Nonmetropolitan · n = 167, median 6.1



4 5 6 7 8 9

Coronary heart disease (% of adults, BRFSS 2023)

Nonmetropolitan minus metropolitan: median difference +0.5; Hodges–Lehmann shift +0.5 (95% CI +0.4 to +0.7) percentage points; probability of superiority 0.73; BH-adjusted $p < 0.0001$ (secondary family).

High blood pressure among adults

Metropolitan · n = 86, median 33.0



Nonmetropolitan · n = 167, median 33.9



27.5 30 32.5 35 37.5 40

High blood pressure (% of adults, BRFSS 2023)

Nonmetropolitan minus metropolitan: median difference +0.9; Hodges–Lehmann shift +1.0 (95% CI +0.5 to +1.6) percentage points; probability of superiority 0.64; BH-adjusted $p = 0.0005$ (secondary family).

Diagnosed diabetes among adults

Metropolitan · n = 86, median 11.6



Nonmetropolitan · n = 167, median 12.5



5 10 15 20 25

Diabetes (% of adults, BRFSS 2023)

Nonmetropolitan minus metropolitan: median difference +0.9; Hodges-Lehmann shift +0.8 (95% CI +0.3 to +1.3) percentage points; probability of superiority 0.61; BH-adjusted p = 0.0033 (secondary family).

Obesity among adults

Metropolitan · n = 86, median 36.4



Nonmetropolitan · n = 167, median 38.0



25 30 35 40 45 50

Obesity (% of adults, BRFSS 2023)

Nonmetropolitan minus metropolitan: median difference +1.6; Hodges-Lehmann shift +1.4 (95% CI +0.6 to +2.2) percentage points; probability of superiority 0.63; BH-adjusted p = 0.0013 (secondary family).

Current cigarette smoking among adults

Metropolitan · n = 86, median 14.1



Nonmetropolitan · n = 167, median 15.9



5 10 15 20 25

Current smoking (% of adults, BRFSS 2023)

Nonmetropolitan minus metropolitan: median difference +1.8; Hodges-Lehmann shift +1.8 (95% CI +1.2 to +2.3) percentage points; probability of superiority 0.73; BH-adjusted p < 0.0001 (secondary family).

No leisure-time physical activity among adults

Metropolitan · n = 86, median 27.8



Nonmetropolitan · n = 167, median 30.5



10 20 30 40 50

No leisure-time activity (% of adults, BRFSS 2023)

Nonmetropolitan minus metropolitan: median difference +2.8; Hodges-Lehmann shift +2.7 (95% CI +1.5 to +3.9) percentage points; probability of superiority 0.66; BH-adjusted p < 0.0001 (secondary family).

3.3 Secondary family: CDC PLACES estimates

All six PLACES age-adjusted estimates were higher in nonmetropolitan counties, with confidence intervals for the shift above zero and adjusted p ≤ 0.0033 (Table 4, Figure 2). The largest shifts in percentage points were for no leisure-time physical activity (+2.7 percentage points (95% CI +1.5 to +3.9)) and current smoking (+1.8 percentage points (95% CI +1.2 to +2.3)); by probability of superiority the clearest separations were coronary heart disease (0.73) and current smoking (0.73).

How often a nonmetropolitan county has the higher value

Probability of superiority for each indicator: 0.5 means no tendency either way; values above 0.5 mean nonmetropolitan counties tend to be higher.

Primary family (observed rates)



Secondary family (CDC PLACES, age-adjusted)

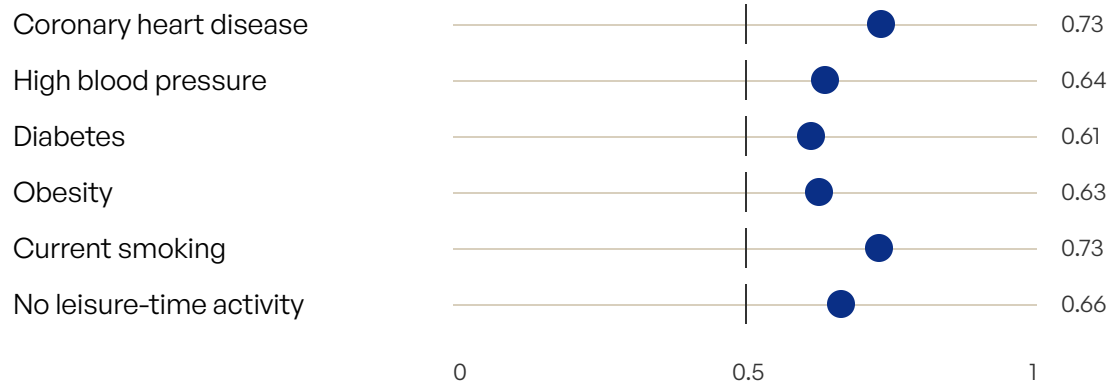


Figure 2. Filled dots: Benjamini–Hochberg-adjusted $p < 0.05$ within the indicator’s family; hollow dots (n.s.): not significant, which does not show the groups are equivalent. The probability of superiority compares counties, not people, and Table 4 gives each shift in the indicator’s own units with its 95% CI.

Table 4. Nonmetropolitan minus metropolitan, by indicator

Indicator	Difference in medians	Hodges–Lehmann shift (95% CI)	Probability of superiority (rank-biserial r)	p	BH-adjusted p
Heart disease deaths (per 100,000) · primary	+53.9	+45.4 (+25.7 to +65.0)	0.67 (0.34)	< 0.0001	< 0.0001
Lung cancer incidence (per 100,000) · primary	+2.1	+3.5 (–0.9 to +7.9)	0.56 (0.13)	0.1340	0.1340
Coronary heart disease (%)	+0.5	+0.5 (+0.4 to +0.7)	0.73 (0.47)	< 0.0001	< 0.0001
High blood pressure (%)	+0.9	+1.0 (+0.5 to +1.6)	0.64 (0.28)	0.0003	0.0005
Diabetes (%)	+0.9	+0.8 (+0.3 to +1.3)	0.61 (0.23)	0.0033	0.0033
Obesity (%)	+1.6	+1.4 (+0.6 to +2.2)	0.63 (0.25)	0.0011	0.0013
Current smoking (%)	+1.8	+1.8 (+1.2 to +2.3)	0.73 (0.46)	< 0.0001	< 0.0001
No leisure-time activity (%)	+2.8	+2.7 (+1.5 to +3.9)	0.66 (0.33)	< 0.0001	< 0.0001

3.4 Missing data

Lung cancer incidence is suppressed when a county had fewer than 16 cases over 2019–2023. That affected 43 of 168 nonmetropolitan counties (26%) and 8 of 86 metropolitan counties (9%), a difference unlikely to be due to chance (Fisher’s exact $p = 0.0026$).

Suppressed counties are small: median 2020 population 2,547 (largest 7,694), against 24,435 for counties with a rate. The nonmetropolitan lung cancer comparison therefore describes mainly the larger nonmetropolitan counties, and its result could differ if the smallest counties could be included. For the other seven indicators only Loving County (population 64) was suppressed.

Table 5. Counties with suppressed lung cancer incidence, by group

Group	Counties
Nonmetropolitan (43)	Bailey, Borden, Briscoe, Collingsworth, Cottle, Crockett, Culberson, Dallam, Dickens, Edwards, Fisher, Floyd, Foard, Glasscock, Hall, Hartley, Hemphill, Jeff Davis, Jim Hogg, Kenedy, Kent, Kimble, King, Kinney, La Salle, Lipscomb, Loving, McMullen, Mason, Menard, Motley, Presidio, Reagan, Roberts, Schleicher, Sherman, Sterling, Stonewall, Sutton, Terrell, Throckmorton, Upton, Yoakum
Metropolitan (8)	Armstrong, Cochran, Garza, Hudspeth, Irion, Lynn, Martin, Oldham

Reason given by State Cancer Profiles: Suppressed by State Cancer Profiles to protect confidentiality and rate stability: fewer than 16 cases were reported over 2019–2023 (average annual count 3 or fewer).

3.5 Exploratory analyses

E1, crude prevalence. With crude rather than age-adjusted PLACES prevalence, every shift remained positive and most were larger (for example coronary heart disease +1.5 against +0.5 percentage points), consistent with nonmetropolitan counties having older adult populations; this is why age-adjusted values are the main analysis.

E2, gradient. Every indicator was positively correlated with the RUCC code (p 0.11 to 0.31), so values tended to rise from the most metropolitan to the least populated, least connected counties; the correlation for diabetes was not statistically significant.

E3, adjacency. Among nonmetropolitan counties, those adjacent to a metropolitan area did not differ clearly from nonadjacent ones on any indicator (all $p > 0.05$; heart disease deaths shift -0.3 , 95% CI -23.9 to $+22.1$ per 100,000).

Table 6. Exploratory analyses (unadjusted p-values)

Indicator	E1 crude shift (95% CI)	E2 ρ with RUCC code (p)	E3 nonadjacent minus adjacent shift (95% CI)	E3 p
Heart disease deaths (per 100,000)	Not applicable	0.26 (< 0.0001)	-0.3 (-23.9 to +22.1)	0.9741
Lung cancer incidence (per 100,000)	Not applicable	0.21 (0.0033)	-5.2 (-11.6 to +1.2)	0.1250
Coronary heart disease (%)	+1.5 (+1.2 to +1.9)	0.31 (< 0.0001)	0.0 (-0.2 to +0.2)	0.9222
High blood pressure (%)	+3.7 (+2.7 to +4.6)	0.14 (0.0274)	-0.3 (-1.0 to +0.3)	0.3542
Diabetes (%)	+2.0 (+1.4 to +2.5)	0.11 (0.0689)	-0.1 (-0.7 to +0.6)	0.8289
Obesity (%)	+1.1 (+0.3 to +1.9)	0.14 (0.0282)	-0.3 (-1.2 to +0.5)	0.4239
Current smoking (%)	+1.5 (+0.9 to +2.0)	0.30 (< 0.0001)	-0.4 (-1.0 to +0.2)	0.1974
No leisure-time activity (%)	+3.5 (+2.4 to +4.7)	0.19 (0.0029)	-0.3 (-1.8 to +1.2)	0.7231

4. Discussion

In current Texas data, nonmetropolitan counties tended to have higher heart disease death rates and higher modelled prevalence of coronary heart disease, high blood pressure, diabetes, obesity, smoking and physical inactivity than metropolitan counties. The direction of the heart disease finding matches national CDC reports.^[3] Because the death rate is age-standardized, the difference is not simply a result of nonmetropolitan counties having older residents, but this study cannot say what does explain it: income, education, insurance, occupation, access to care and many other factors differ between the groups and are not analysed here.

The lung cancer comparison did not show a clear difference, unlike national findings of higher incidence in nonmetropolitan counties. ^[4] It should be read as inconclusive rather than as evidence of no difference: the confidence interval is wide, and a quarter of nonmetropolitan counties had no published rate, so the comparison mostly describes the larger nonmetropolitan counties. The PLACES differences are consistent with the death-rate finding but are weaker evidence on their own, because the estimates come from a model that uses county demographics.

The differences are between distributions of counties. A probability of superiority of 0.67 for heart disease deaths means that the groups overlap substantially: many nonmetropolitan counties have lower rates than many metropolitan ones. Within the nonmetropolitan group, adjacency to a metropolitan area made no clear difference.

5. Limitations

- **Ecological fallacy and no individual data.** Every value describes a county population. Differences between counties cannot be applied to the people living in them, and no individual records were used. ^[15, 16]
- **Confounding.** Metropolitan status travels with age structure beyond what adjustment removes, income, education, insurance, occupation, environment and access to care. None is adjusted for, and the comparisons do not show that metropolitan status causes any difference.
- **County-level classification.** RUCC classifies whole counties. Metropolitan counties include rural areas and nonmetropolitan counties include towns, so the categories mask large differences within counties. ^[2]
- **Different reporting periods.** Deaths cover 2022–2024, cancer diagnoses 2019–2023, PLACES estimates 2023 and the classification 2023.
- **Model-based estimates.** PLACES values are small-area model estimates that use county demographics; differences between groups partly reflect the model's inputs. ^[9]
- **Suppression and selection.** 51 small counties, 43 of them nonmetropolitan, have no published lung cancer rate, so that comparison describes larger counties. The heart disease death rate is spatially smoothed, which borrows strength from neighbouring counties and can reduce differences between them. ^[6]

- **Population differences.** Counties are unweighted. Metropolitan counties held 90% of Texans, so the results describe counties, not the average resident of either group.
- **Age adjustment.** The three sources adjust to different populations (adults 35+, all ages, adults 18+), and PLACES does not document its standard population, so indicators are never compared with each other.
- **Uncertainty and multiple comparisons.** County estimates carry their own uncertainty, which these tests do not include, and neighbouring counties are not independent, which can make p-values look more certain than they are. Eight comparisons were tested; adjustment controls the false discovery rate within each family but not across the exploratory analyses.
- **Not peer reviewed.** This is an independent analysis, checked by an independent recomputation but not by external reviewers.

6. Conclusion

Among Texas counties, nonmetropolitan counties tended to have higher heart disease death rates (Hodges–Lehmann shift +45.4 per 100,000 (95% CI +25.7 to +65.0)) and higher modelled prevalence of six cardiometabolic conditions and risk behaviours than metropolitan counties. Lung cancer incidence did not differ clearly, a result limited by suppression in small nonmetropolitan counties. These are ecological, cross-sectional differences between counties; they do not show that metropolitan status causes them and do not describe any individual’s health. This report is not medical advice.

AI assistance. Anthropic’s Claude, used through Claude Code, produced most of Vellora’s software and research materials under the author’s direction: the site’s code and architecture; the scripts that download, validate and process the source data; the statistical analysis code and the separate Python verification scripts; the charts and maps; the PDF generation; the analysis plans; the drafting and editing of the report text; and the documentation and reproducibility package. No other AI tools were used. Computational results were checked by reproducible scripts, including an independent Python re-implementation of each study’s statistics, which was itself AI-written. These checks show that the reported numbers follow from the code and data; they are not independent scientific peer review. Claude is a tool, not an author.

Agamvir Singh set the research questions and requirements, has read and reviewed each published report, and is responsible for reviewing the published work and its interpretations. The research is independent and exploratory and has not been peer reviewed.

Independent exploratory research — not peer reviewed. This report is by Agamvir Singh for Vellora, is not affiliated with any university or public agency, and is not medical advice. USDA, CDC, NCI and Texas Cancer Registry data are used without endorsement by any of them.

7. References

- [1] U.S. Department of Agriculture, Economic Research Service. Rural-Urban Continuum Codes, 2023 (data and documentation). Accessed 2026-10-23. <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes>
- [2] U.S. Department of Agriculture, Economic Research Service. Rural Classifications: What is Rural? Accessed 2026-10-23. <https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/what-is-rural>
- [3] Moy E, Garcia MC, Bastian B, et al. Leading causes of death in nonmetropolitan and metropolitan areas — United States, 1999–2014. MMWR Surveill Summ. 2017;66(1):1–8. <https://doi.org/10.15585/mmwr.ss6601a1>
- [4] Henley SJ, Anderson RN, Thomas CC, Massetti GM, Peaker B, Richardson LC. Invasive cancer incidence, 2004–2013, and deaths, 2006–2015, in nonmetropolitan and metropolitan counties — United States. MMWR Surveill Summ. 2017;66(14):1–13. <https://doi.org/10.15585/mmwr.ss6614a1>
- [5] Centers for Disease Control and Prevention, Division for Heart Disease and Stroke Prevention. Heart Disease Mortality Data Among US Adults (35+) by State/Territory and County – 2022–2024 (dataset mri7-5jtw). Accessed 2026-10-18. <https://data.cdc.gov/d/mri7-5jtw>
- [6] Centers for Disease Control and Prevention. Atlas of Heart Disease and Stroke: Statistical Methods. Accessed 2026-10-22. <https://www.cdc.gov/heart-disease-stroke-atlas/statistical-methods/index.html>
- [7] National Cancer Institute and Centers for Disease Control and Prevention. State Cancer Profiles: Incidence Rate Report for Texas by County, Lung & Bronchus (All Stages), 2019–2023. Data: Texas Cancer Registry, SEER November 2025 submission. Accessed 2026-10-18. <https://statecancerprofiles.cancer.gov/incidencerates/index.php?stateFIPS=48&areatype=county&cancer=047&race=00&sex=0&age=001&stage=999&year=0&type=incd&sortVariableName=rate&sortOrder=default>

- [8] Centers for Disease Control and Prevention. PLACES: Local Data for Better Health, County Data, 2025 release (dataset swc5-untb; rows updated 2025-12-04). Accessed 2026-10-22. <https://data.cdc.gov/d/swc5-untb>
- [9] Centers for Disease Control and Prevention. PLACES: Methodology. Accessed 2026-10-23. <https://www.cdc.gov/places/methodology/index.html>
- [10] Centers for Disease Control and Prevention. PLACES Data FAQs. Last reviewed 2024-08-29. Accessed 2026-10-23. <https://www.cdc.gov/places/faqs/data-faqs/index.html>
- [11] Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat. 1947;18(1):50–60. <https://doi.org/10.1214/aoms/1177730491>
- [12] Hodges JL, Lehmann EL. Estimates of location based on rank tests. Ann Math Stat. 1963;34(2):598–611. <https://doi.org/10.1214/aoms/1177704172>
- [13] Bauer DF. Constructing confidence sets using rank statistics. J Am Stat Assoc. 1972;67(339):687–690. <https://doi.org/10.1080/01621459.1972.10481279>
- [14] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B. 1995;57(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- [15] Robinson WS. Ecological correlations and the behavior of individuals. Am Sociol Rev. 1950;15(3):351–357. <https://doi.org/10.2307/2087176>
- [16] Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Annu Rev Public Health. 1995;16:61–81. <https://doi.org/10.1146/annurev.pu.16.050195.000425>

8. Reproducibility

Every number in this report, its PDF and its figures is read from one results file, generated by the analysis code from Vellora’s verified copies of the source files and following the analysis plan, which was committed before any group difference was computed. The analysis is deterministic. An independent Python implementation using SciPy and NumPy recomputed every group size, descriptive statistic, Hodges–Lehmann interval, Mann–Whitney test, adjusted p-value, Fisher test and exploratory result, and matched the results file. Where the measures overlap, the group medians and Mann–Whitney tests also match the earlier cardiovascular report exactly. Related: [Smoking Prevalence and Lung Cancer Incidence Across Texas Counties](#).

Analysis plan (Markdown)

Indicators, families, methods and sensitivity analyses, committed before any group difference was computed

Results (JSON)

Every statistic reported here, the suppressed counties with their reasons, and the per-county values in Figure 1

Analysis code (TypeScript)

The analysis that produced the results file

Independent check (Python)

Recomputes every statistic with SciPy and NumPy

PLACES inputs (JSON)

Vellora's verified copy of CDC PLACES 2025 age-adjusted estimates; crude values are in /data/places-tx.v2025.1.json

Heart disease mortality input (JSON)

Vellora's verified copy of CDC county heart disease death rates

Lung cancer input (JSON)

Vellora's verified copy of State Cancer Profiles county incidence

County classification input (JSON)

USDA ERS RUCC 2023 codes and 2020 Census populations

To reproduce: regenerate the results with `npm run research:rural` and check them with `uvx --with scipy --with numpy python scripts/verify-rural-study.py`. The input files record the source releases and download dates; the sources may revise their data.