Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Oral-Brain Axis and Neighborhood Deprivation: An Ecological Data-Warehouse Analysis of Tooth Loss, Cognitive Disability, and Mental Distress in U.S. ZIP Code Areas

Mandvi Thoma

dal062164@utdallas.edu

The University of Texas at Dallas https://orcid.org/0009-0003-2267-0136

Research Article

Keywords: oral microbiome, dysbiosis, periodontitis, tooth loss, cognitive disability, Area Deprivation Index, CDC PLACES, data warehouse, geospatial analysis, health disparities

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11600131/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background. Oral microbial dysbiosis and periodontal disease have been linked to systemic inflammation and cognitive decline, an idea often called the oral-brain axis. Oral disease and cognitive impairment both vary with social conditions, but few analytic systems link neighborhood deprivation to oral and cognitive health across the whole country.

Methods. We built a staged SQL data warehouse that combines the 2024 Area Deprivation Index (ADI; 242,297 block group records keyed by 12-digit FIPS and GISJOIN) with the CDC PLACES extract (1,171,563 rows). For four indicators (complete tooth loss in adults aged 65 and older, cognitive disability, frequent mental distress, and dental visit in the past year), ZIP Code Tabulation Area (ZCTA) prevalence was linked to county-mean ADI. We estimated Spearman correlations, ADI-quintile gradients, and population-weighted regressions with county-clustered standard errors. Microbiome data were not analyzed.

Results. In 29,903 ZCTAs with all four indicators, tooth loss correlated strongly with cognitive disability ($\rho = 0.83$) and mental distress ($\rho = 0.72$), and inversely with dental visits ($\rho = -0.84$). From the least to the most deprived ADI quintile, tooth loss rose from 10.5% to 20.2% (ratio 1.93), cognitive disability from 13.7% to 18.2%, and mental distress from 16.0% to 18.7%, while dental visits fell from 64.4% to 50.4%. Tooth loss was associated with cognitive disability (standardized $\beta = 0.94$, 95% CI 0.91-0.96); this was unchanged after adjustment for ADI ($\beta = 0.96$), and ADI had no independent association ($\beta = -0.04$, 95% CI -0.14 to 0.06).

Conclusions. Oral disease, cognitive disability and mental distress tend to occur together in U.S. ZIP areas, and all rise steeply with neighborhood deprivation. These ecological, model-based findings fit a pathway from deprivation through oral disease to cognitive and mental health problems, but they cannot establish one. The pipeline is reproducible and could later be linked to individual-level microbiome cohorts.

1. Introduction

The mouth carries one of the most diverse microbial communities in the human body. In health, this community is relatively stable; in periodontal disease, it shifts toward a pathogen-enriched, pro-inflammatory configuration (dysbiosis) that sustains chronic inflammation of tooth-supporting tissues and elevates circulating mediators such as C-reactive protein, interleukin-1β, interleukin-6 and tumor necrosis factor-α [1,2]. Tooth loss, the cumulative end point of periodontal disease and caries, is widely used in surveillance as a marker of lifetime oral disease burden.

The oral-brain axis hypothesis proposes that this oral inflammatory and microbial burden contributes to neuroinflammation and cognitive decline. Evidence for it includes associations of periodontitis and tooth loss with cognitive impairment and dementia in cohort studies and meta-analyses [3,4,5]; detection of *Porphyromonas gingivalis* DNA and its gingipain proteases in post-mortem Alzheimer's disease brain

tissue [6]; and several proposed mechanisms: systemic cytokines that prime microglia, bacterial products crossing a weakened blood-brain barrier, and neural routes [2,7]. The evidence is not conclusive. Reverse causation (cognitive impairment reducing oral self-care), shared risk factors such as diabetes and smoking, and inconsistent results of pathogen-targeted interventions all caution against causal interpretation.

Oral health and cognitive health also differ from place to place. Dental care, insurance, nutrition, housing and chronic stress are all worse in disadvantaged neighborhoods. The Area Deprivation Index (ADI), which ranks census block groups using American Community Survey indicators of income, education, employment and housing, has been associated with a range of health outcomes [8-10]. If deprivation shapes oral disease and cognitive burden together, then place-based surveillance data should show a shared social gradient and strong co-occurrence of oral and cognitive indicators.

Testing this across the whole country is as much a data-engineering problem as a statistical one. The relevant sources use different geographies (block groups, tracts, counties, ZIP areas), identifier conventions (12-digit FIPS, NHGIS GISJOIN), suppression codes and release years, and silent key errors are easy to make. This study has three aims: (1) to engineer a reproducible warehouse pipeline that joins ADI and CDC PLACES with explicit quality controls; (2) to quantify the co-occurrence of tooth loss, dental care, cognitive disability and frequent mental distress across U.S. ZIP areas; and (3) to estimate how these indicators vary with neighborhood deprivation, and whether deprivation retains an association with cognitive and mental health indicators once oral disease burden is considered.

2. Methods

2.1 Study design and data sources

We conducted a cross-sectional, ecological analysis of two public, aggregate datasets. **ADI 2024:** the national ADI file [17] contains 242,297 data records (242,298 lines including the header), one per census block group, with the NHGIS GISJOIN identifier [18], the 12-digit block group FIPS code, the national percentile rank (1-100; higher is more deprived) and the within-state decile (1-10). Records for which no rank is computed carry suppression codes (GQ, group quarters; PH, low population or housing; GQ-PH; QDI, quality-of-data indicator). **CDC PLACES:** the extract contains 1,171,563 rows of model-based small-area estimates of crude prevalence, covering 2022 and 2023 release years, with 5-digit location identifiers that correspond to ZIP Code Tabulation Areas (ZCTAs; for example, 01001 corresponds to Agawam, Massachusetts, consistent with the file's coordinates). PLACES estimates are derived from the Behavioral Risk Factor Surveillance System using multilevel regression and poststratification [11,16].

2.2 Indicators

We analyzed four PLACES measures: complete tooth loss among adults aged 65 years and older (TEETHLOST, release 2022), visit to a dentist or dental clinic in the past year among adults (DENTAL, 2022), cognitive disability among adults (COGNITION, 2023; self-reported serious difficulty

concentrating, remembering or making decisions), and frequent mental distress among adults (MHLTH, 2023; 14 or more days of poor mental health in the past 30 days). Cognitive disability is a self-reported functional measure and not a diagnosis of dementia. Microbiome data are not contained in these sources.

2.3 Data warehouse and geospatial engineering

The pipeline (Figure 1) follows a three-layer design. In the staging layer, files were loaded unchanged with every identifier typed as text, because integer typing silently removes the leading zero of FIPS codes for states 01-09 and breaks joins. The integration layer standardized block group FIPS to 12 characters (state 2 + county 3 + tract 6 + block group 1), derived state, county and tract keys by substring, retained GISJOIN as a parallel key, converted ADI ranks to numeric and recoded suppression codes to missing. The analytical layer holds a star schema [15] (geography and measure dimensions; ADI and PLACES fact tables) from which an analysis mart was produced. Data-quality contracts were evaluated at each load: uniqueness of block group FIPS, key length, value domain (ADI rank 1-100), expected row counts, and match rates between sources. A Tableau-ready extract of the final mart is provided as Supplementary Data S1.

2.4 Linkage of ADI to PLACES

Because PLACES is published at ZCTA level and ADI at block group level, the two sources share no common geography. We aggregated ADI to the county (first five digits of the 12-digit FIPS) as the unweighted mean national rank of non-suppressed block groups (3,222 counties), and assigned each ZCTA to a single county using USPS ZIP-code records, mapping county name and state to county FIPS through a GeoNames lookup. Each ZCTA therefore received the mean ADI of its assigned county. This is a coarse bridge: ZIP areas can cross county lines, county averages mask within-county variation (the modifiable areal unit problem [14]), and the ADI file contains no population field for weighting block groups. We treat the results as a first-pass national screen and recommend a block-group-to-ZCTA areal crosswalk for confirmatory work.

2.5 Statistical analysis

We described each indicator (mean, standard deviation, median) and computed Spearman rank correlations among the four indicators and between each indicator and ADI. ZCTAs were divided into quintiles of ADI, and population-weighted mean prevalence was reported for each quintile together with the ratio of the most to least deprived quintile. We then fitted weighted least-squares regressions of standardized (z-scored) outcomes on standardized predictors, weighting by ZCTA total population and computing standard errors clustered by county to account for ZCTAs sharing a county ADI value. Models examined (i) ADI alone for each indicator; (ii) cognitive disability and mental distress on tooth loss, without and with ADI; and (iii) cognitive disability on dental visit, without and with ADI. Regression models used ZCTAs with all four indicators and ADI available (n = 29,588 in 2,941 counties). Analyses were performed in Python (pandas [19], statsmodels [20], SciPy [21]). Two-sided 95% confidence

intervals are reported; because the samples are very large, we focus on effect sizes rather than p-values [22].

2.6 Ethics

This study used publicly available, aggregate, de-identified data and did not involve human participants; institutional review was not required.

2.7 Use of generative AI

Generative AI (Claude, Anthropic) was used during manuscript preparation to help draft and edit the text and to assist with the analysis code. The author specified the research question, obtained the data sources, ran and checked the analyses, reviewed and verified all AI-assisted text and code, and takes full responsibility for the content of the manuscript.

3. Results

3.1 Data quality and linkage

Of the 242,297 ADI records, 6,132 (2.5%) carried suppression codes (GQ, 2,595; PH, 2,522; GQ-PH, 789; QDI, 226), leaving 236,165 block groups with a national rank across 3,222 counties. Every 12-digit FIPS was unique and of correct length. The PLACES extract included 32,520 ZCTAs for dental visit, 32,437 for tooth loss, and 29,983 for each of cognitive disability and mental distress, with no duplicate ZCTA-measure combinations. All ZCTAs were found in the ZIP-record lookup, 99.9% were assigned a county FIPS, and 32,204 received an ADI value; 29,903 had all four indicators. Table 1 summarizes the indicators.

Table 1. Descriptive statistics of PLACES indicators across ZIP Code Tabulation Areas.

Indicator	Release	ZCTAs (n)	Mean (%)	SD	Median (%)
Complete tooth loss, age ≥65	2022	32,437	14.7	6.7	13.5
Dental visit, past year	2022	32,520	60.7	10.3	62.2
Cognitive disability	2023	29,983	15.2	3.6	14.8
Frequent mental distress	2023	29,983	17.0	2.7	16.9

Crude prevalence from CDC PLACES. SD, standard deviation across ZCTAs (unweighted).

3.2 Co-occurrence of oral, cognitive and mental health indicators

Indicators were strongly correlated (Table 2; Figure 2). Tooth loss was positively correlated with cognitive disability ($\rho = 0.83$) and mental distress ($\rho = 0.72$) and negatively correlated with dental visits

(ρ = -0.84). Dental visits were inversely correlated with cognitive disability (ρ = -0.85) and mental distress (ρ = -0.73). Cognitive disability and mental distress were themselves very closely related (ρ = 0.90).

Table 2. Spearman rank correlations among indicators (n = 29,903 ZCTAs with all four).

	Tooth loss	Cognitive disability	Mental distress	Dental visit
Tooth loss	1.00	0.83	0.72	-0.84
Cognitive disability	0.83	1.00	0.90	-0.85
Mental distress	0.72	0.90	1.00	-0.73
Dental visit	-0.84	-0.85	-0.73	1.00

All correlations p < 0.001.

3.3 Deprivation gradient

All four indicators varied monotonically across ADI quintiles (Figure 3; Table 3). Between the least and most deprived quintiles, population-weighted tooth loss rose from 10.5% to 20.2% (ratio 1.93), cognitive disability from 13.7% to 18.2% (1.33), and mental distress from 16.0% to 18.7% (1.17), while dental visits fell from 64.4% to 50.4% (0.78). Rank correlations with ADI were 0.55 for tooth loss, 0.42 for cognitive disability, 0.36 for mental distress and -0.46 for dental visit (n = 32,122; 29,667; 29,667; 32,204, respectively). The steepest gradient was therefore for tooth loss.

Table 3. Prevalence by ADI quintile and rank correlation with ADI.

Indicator (%)	Q1	Q2	Q3	Q4	Q5	Q5/Q1
Tooth loss, age ≥65	10.5	12.6	14.8	16.7	20.2	1.93
Cognitive disability	13.7	14.6	15.7	16.5	18.2	1.33
Mental distress	16.0	16.8	17.6	18.3	18.7	1.17
Dental visit	64.4	61.0	60.0	57.7	50.4	0.78

Population-weighted means by quintile of county-mean ADI national rank (about 6,430 ZCTAs per quintile).

3.4 Regression analyses

In models with a single predictor, a one-standard-deviation higher ADI was associated with higher tooth loss (β = 0.41, 95% CI 0.37-0.46; R^2 = 0.22), lower dental visits (β = -0.33, -0.41 to -0.25; R^2 = 0.12), higher cognitive disability (β = 0.36, 0.26-0.45; R^2 = 0.13) and higher mental distress (β = 0.35, 0.29-0.41; R^2 = 0.14). Tooth loss was strongly associated with cognitive disability (β = 0.94, 0.91-0.96; R^2 = 0.72) and with mental distress (β = 0.80, 0.76-0.83; R^2 = 0.57). Adding ADI left these estimates essentially

unchanged (cognitive disability: tooth loss β = 0.96, 0.90-1.02, ADI β = -0.04, -0.14 to 0.06; mental distress: tooth loss β = 0.78, 0.73-0.83, ADI β = 0.03, -0.04 to 0.09). Dental visit was inversely associated with cognitive disability (β = -0.80, -0.87 to -0.73), attenuating modestly with ADI (β = -0.76, -0.84 to -0.68), with a small independent ADI association (β = 0.11, 0.04-0.17).

Thus, in ZIP areas with a similar level of tooth loss, deprivation carried no additional association with cognitive disability or mental distress; the deprivation gradient in these outcomes was largely shared with the gradient in oral disease burden.

4. Discussion

Across more than 29,000 ZIP areas, tooth loss, infrequent dental care, cognitive disability and frequent mental distress occurred together, and each rose or fell steadily with neighborhood deprivation. Tooth loss showed the steepest gradient, almost doubling from the least to the most deprived quintile. After accounting for tooth loss, ADI had no independent association with cognitive disability or mental distress. This pattern is what a pathway from deprivation to oral disease burden to cognitive and mental health burden would predict, and it is also what shared upstream determinants, common model inputs, or confounding by age, diabetes, smoking and insurance would produce. The data cannot distinguish these explanations.

Cognitive disability and mental distress are even more strongly correlated with each other (ρ = 0.90) than with the oral indicators, which calls for caution. Both are self-reported and share modeling predictors in PLACES, and depression, cognition and oral self-care interact at the individual level. Any oral-brain interpretation therefore has to be separated from a general “health-burden” factor that includes cognition, mood and care access. Mechanistic studies of oral microbial profiles, inflammatory markers and neuroimaging in individual-level cohorts are required for that separation [3,6,7].

From a public-health standpoint, the large gradients in tooth loss and dental visits suggest access barriers that could be changed (dental insurance, workforce distribution, fluoridation, preventive services) concentrated in deprived areas. Oral health is rarely part of dementia-prevention and mental health programs. Our results support piloting that kind of integration in high-deprivation areas and monitoring oral indicators as a component of healthy-aging surveillance.

On the informatics side, the study shows how to avoid common silent errors in geospatial health data: text-typed FIPS identifiers, explicit suppression handling, contract tests on keys and counts, and a documented, versioned linkage rule. The mismatch between source geographies (block group versus ZCTA) was the biggest engineering problem, and anyone starting an ADI-PLACES project should check it early.

4.1 Limitations

- **Ecological design.** Associations between areas do not describe individuals [13]; the people with tooth loss in a ZIP area are not necessarily those with cognitive disability.
- **Coarse geographic linkage.** County-mean ADI assigned to ZIP areas ignores within-county variation and ZIP-county overlap, and is unweighted by population. This likely attenuates true associations with ADI and may distort their shape.
- **Model-based outcomes.** PLACES estimates are smoothed by shared covariates, which inflates correlations among measures and with demographic and deprivation variables.
- **Construct validity.** Cognitive disability is not dementia; tooth loss and dental visits are proxies for oral disease and care, not measures of periodontitis or microbial dysbiosis. No microbiome data were analyzed.
- **Temporal mismatch.** Indicators come from two release years and ADI from a different survey period; the cross-sectional design cannot establish temporality or causation.
- **Residual confounding and unadjusted covariates.** Age structure, race and ethnicity, diabetes, smoking and insurance were not adjusted for, and spatial autocorrelation was not modeled; standard errors clustered by county may still be optimistic.

4.2 Future work

Next steps include block-group-to-ZCTA or tract-level linkage with population weights; covariate-adjusted multilevel and spatial models (spatial lag or error, Bayesian conditional autoregressive models) with tests for residual autocorrelation; local indicators of spatial association to map clusters of concurrent deprivation, tooth loss and cognitive disability [12]; and linkage to individual-level cohorts with oral microbiome sequencing, cognitive testing and residential geocodes.

5. Conclusion

Tooth loss, limited dental care, cognitive disability and frequent mental distress co-occur across U.S. ZIP areas and share a steep gradient with neighborhood deprivation, and the deprivation signal is largely absorbed by oral disease burden. These population-level patterns give a reason to study the oral-brain axis hypothesis but do not test it. The warehouse pipeline and a Tableau-ready mart are provided so that others can run finer spatial analyses and later link to microbiome cohorts.

Declarations

Ethics approval and consent to participate: Not applicable; analysis of public aggregate data.

Consent for publication: Not applicable.

Data availability: The ADI 2024 data are available from the Neighborhood Atlas (University of Wisconsin School of Medicine and Public Health) and CDC PLACES data from the Centers for Disease Control and

Prevention. The linked ZCTA-level analysis mart is provided as Supplementary Data S1; analysis code is available from the corresponding author upon reasonable request.

Funding: This research received no external funding.

Author contributions (CRediT): Conceptualization, methodology, software, data curation, formal analysis, visualization, writing (original draft preparation), writing (review and editing), and supervision, M.T. The author has read and approved the final version of the manuscript.

Acknowledgments: Area Deprivation Index data were obtained from the Neighborhood Atlas, University of Wisconsin School of Medicine and Public Health.

Competing interests: The author declares no competing interests. Neither the author nor the author's institution received payments or services from a third party in the past 36 months that could be perceived to influence the submitted work.

References

1. Hajishengallis G. Periodontitis: from microbial immune subversion to systemic inflammation. *Nat Rev Immunol*. 2015;15(1):30-44.
2. Hajishengallis G, Chavakis T. Local and systemic mechanisms linking periodontal disease and inflammatory comorbidities. *Nat Rev Immunol*. 2021;21(7):426-440.
3. Kamer AR, Craig RG, Dasanayake AP, Brys M, Glodzik-Sobanska L, de Leon MJ. Inflammation and Alzheimer's disease: possible role of periodontal diseases. *Alzheimers Dement*. 2008;4(4):242-250.
4. Nadim R, Tang J, Dilmohamed A, et al. Influence of periodontal disease on risk of dementia: a systematic literature review and a meta-analysis. *Eur J Epidemiol*. 2020;35(9):821-833.
5. Ide M, Harris M, Stevens A, et al. Periodontitis and cognitive decline in Alzheimer's disease. *PLoS One*. 2016;11(3):e0151081.
6. Dominy SS, Lynch C, Ermini F, et al. *Porphyromonas gingivalis* in Alzheimer's disease brains: evidence for disease causation and treatment with small-molecule inhibitors. *Sci Adv*. 2019;5(1):eaau3333.
7. Kamer AR, Pushalkar S, Gulivindala D, et al. Periodontal dysbiosis associates with reduced CSF Aβ42 in cognitively normal elderly. *Alzheimers Dement (Amst)*. 2021;13(1):e12172.
8. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible: the Neighborhood Atlas. *N Engl J Med*. 2018;378(26):2456-2458. doi:10.1056/NEJMp1802313
9. Kind AJH, Jencks S, Brock J, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. *Ann Intern Med*. 2014;161(11):765-774.
10. Singh GK. Area deprivation and widening inequalities in US mortality, 1969-1998. *Am J Public Health*. 2003;93(7):1137-1143.
11. Zhang X, Holt JB, Lu H, et al. Multilevel regression and poststratification for small-area estimation of population health outcomes: a case study of chronic obstructive pulmonary disease prevalence

using the BRFSS. *Am J Epidemiol*. 2014;179(8):1025-1033.

12. Anselin L. Local indicators of spatial association, LISA. *Geogr Anal*. 1995;27(2):93-115.
13. Robinson WS. Ecological correlations and the behavior of individuals. *Am Sociol Rev*. 1950;15(3):351-357.
14. Openshaw S. *The Modifiable Areal Unit Problem*. Norwich: Geo Books; 1984.
15. Kimball R, Ross M. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. 3rd ed. Indianapolis: Wiley; 2013.
16. Centers for Disease Control and Prevention. PLACES: Local Data for Better Health. Methodology and data dictionary.https://www.cdc.gov/places.
17. University of Wisconsin School of Medicine and Public Health. 2024 Area Deprivation Index. Downloaded from https://www.neighborhoodatlas.medicine.wisc.edu.
18. IPUMS NHGIS. GISJOIN identifiers and geographic crosswalks. https://www.nhgis.org
19. McKinney W. Data structures for statistical computing in Python. *Proc 9th Python in Science Conf*. 2010:56-61.
20. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with Python. *Proc 9th Python in Science Conf*. 2010:92-96.
21. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17:261-272.
22. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. *Am Stat*. 2016;70(2):129-133.

Figures

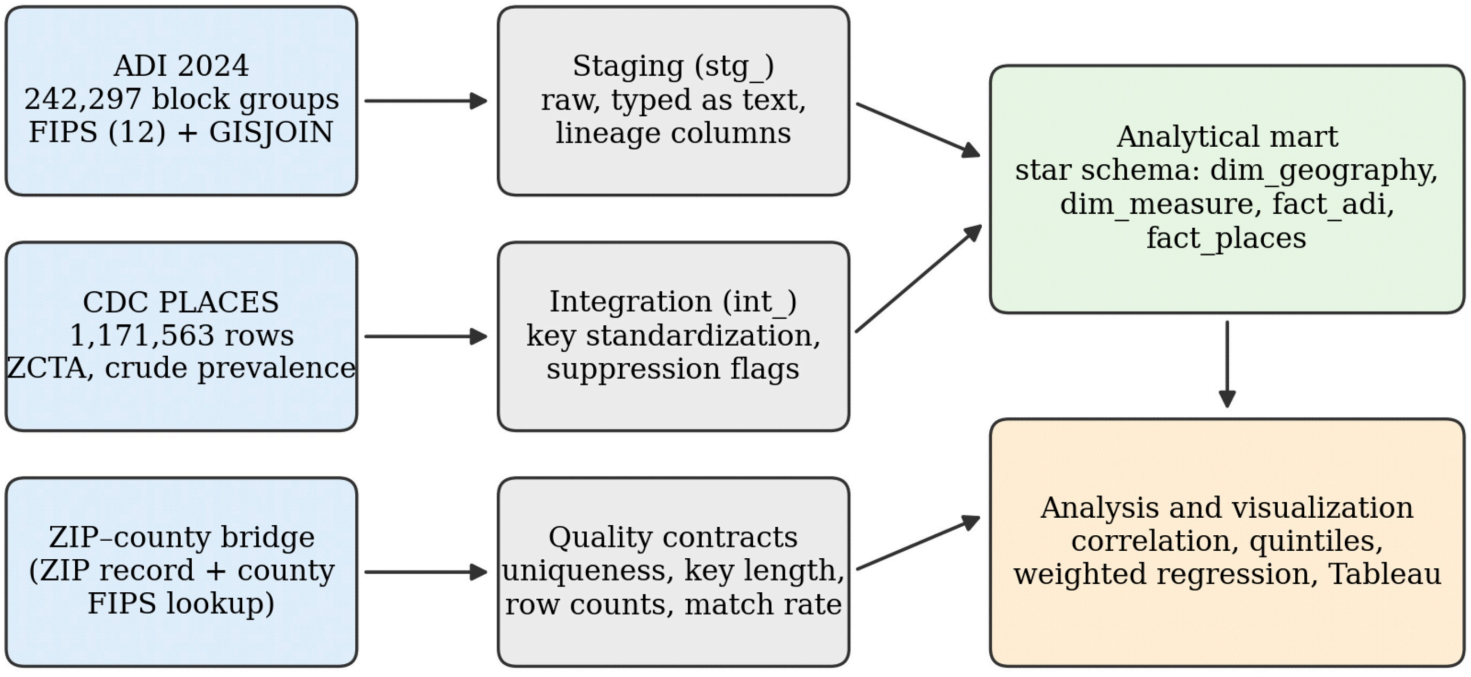


Figure 1

Data-warehouse and linkage pipeline. Sources are staged as text, standardized and checked against quality contracts, and combined in an analytical mart.

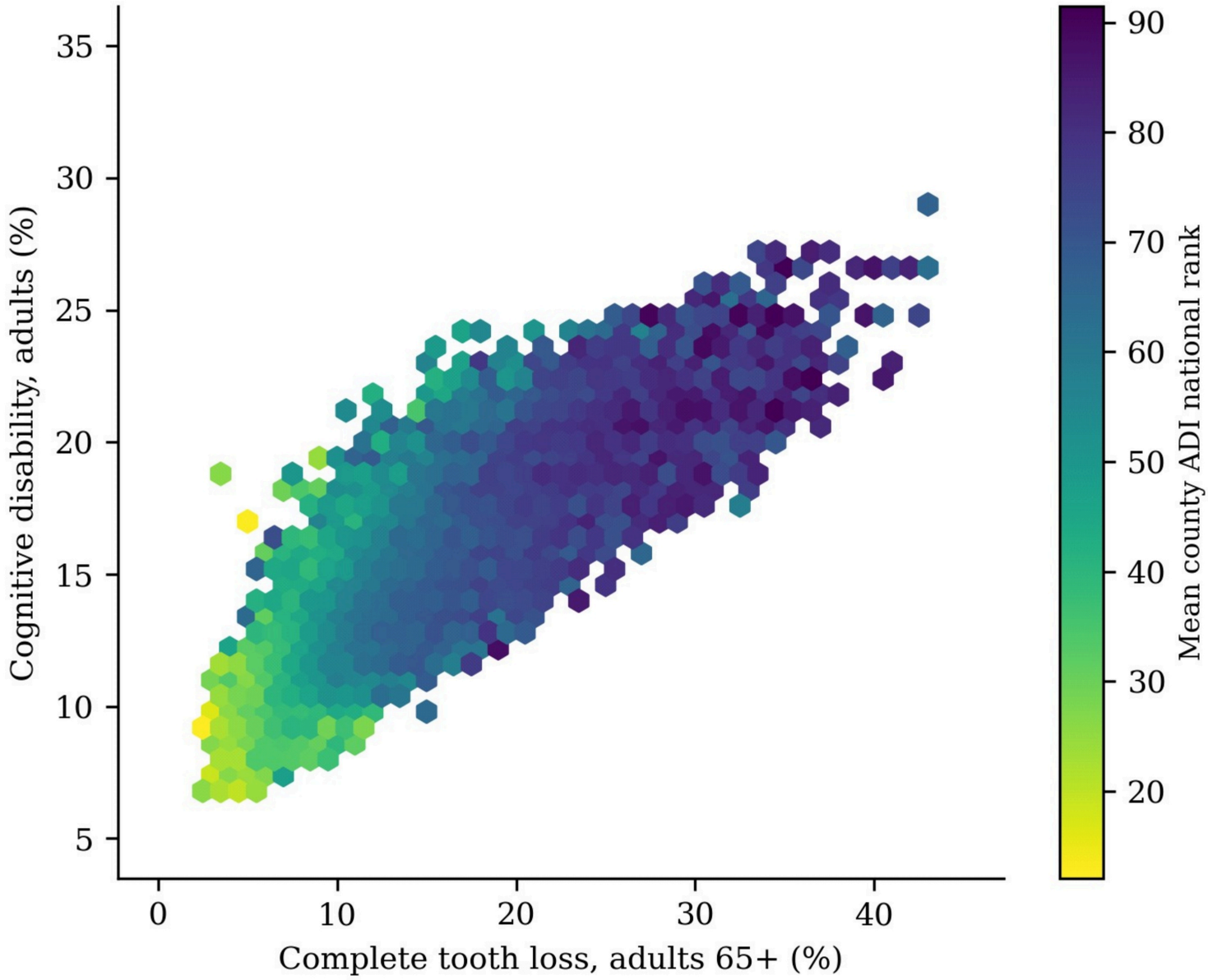


Figure 2

Relationship between tooth loss and cognitive disability across ZCTAs; color shows the mean ADI national rank of ZCTAs in each hexagonal bin (bins with fewer than three ZCTAs omitted).

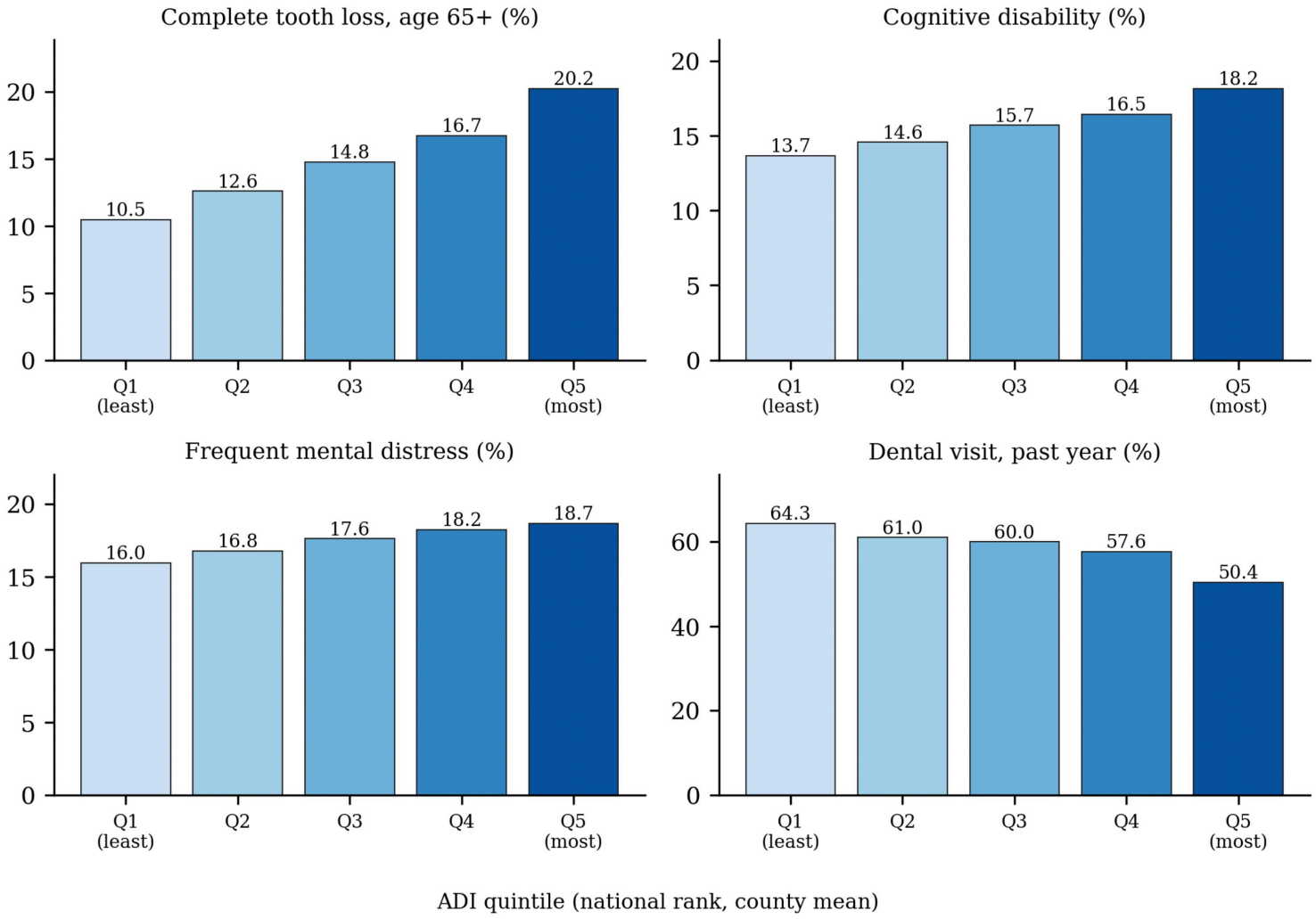


Figure 3

Population-weighted mean prevalence of four indicators by ADI quintile (Q1, least deprived; Q5, most deprived).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryDataS1zctaadimart.csv