

TITLE: A century of census tracts: health & the body politic (1906-2006)

AUTHOR: Nancy Krieger, PhD
Professor, Department of Society, Human Development and Health
Harvard School of Public Health
677 Huntington Avenue
Boston, MA, USA 02115

office: 617-432-1571
fax: 617-432-3123
email: nkrieger@hsph.harvard.edu

WORD COUNT: 1997 words (text only, not including abstract or references)

FIGURES: 1

Acknowledgements: Thanks to Pamela D. Waterman for generating the Year 2000 New York City census tract map. Preparation of this manuscript was not supported by any grant.

Submitted to *Journal of Urban Health* on November 1, 2005

Revised and resubmitted on December 16, 2005

Accepted December 19, 2005

ABSTRACT

In 2006, the US will celebrate the 100th birthday of the census tract. These geographic units, born out of concerns for urban well-being, were first proposed in 1906 to provide a “convenient and scientific city map system” for the City of New York. They were employed for the first time in the US census in 1910 in 8 cities, via a joint effort involving the US Census Bureau and state and local health departments.

Initially termed “sanitary areas” because of their relevance to planning for public health and health services, census tracts are now widely used by all sectors of government and by myriad disciplines in the health, social, and geographic sciences for research as well as policy development, implementation, and evaluation. In this article, I describe the census tract’s underappreciated origins, give examples of its current use in analyzing and addressing social disparities in health and health care, and discuss its continued significance and implications for population health and the public data required for informed democratic governance.

KEY WORDS: census tract, geocoding, geography, GIS, health disparities, policy, public health surveillance, socioeconomic, urban health.

In 2006, the United States census tract will celebrate its 100th birthday. This remarkable – but often taken for granted – parcel of government-defined geography occupies an important place in our body politic, with direct relevance for addressing inequities in health and health care. Born at the crossroads of planning for congregation size and public health, census tracts are now widely used by all sectors of government and by myriad academic disciplines to describe, analyze, and improve societal well-being. A centenary offers a useful moment to consider the census tract’s underappreciated origins and its significance and implications for population health and the public data required for informed democratic governance.

History: from “sanitary areas” to “census tracts”

In brief, the idea of establishing small permanent geographic areas was first proposed in 1906 by Dr. Walter Laidlaw, then Director of the Population Research Bureau of the New York Federation of Churches (Figure 1).¹ Stymied by the then conventional practice of enumerating populations by political wards – whose boundaries changed with each election – Laidlaw’s “strong conviction was that unanticipated population movements and general ignorance of neighborhoods and neighbors was gravely hurting church and synagogue efficiency,”^{2, p. 4} since they could neither track nor anticipate changes in their congregation base. Intent on comparing areas’ population counts over time, using stable boundaries, Laidlaw succeeded in convincing the Chief of the Division of Population of the Census Bureau of his idea, and in 1910 the Census Bureau created the first ever small census districts for New York City and 7 other large cities (Figure 1). Initially, these districts were called “sanitary areas” because of health departments’ role in establishing their boundaries, given the areas’ relevance to planning for public health and health services, e.g., by enabling calculation and tracking of neighborhood morbidity and mortality rates. By the mid-1930s, however, as their use by non-health agencies and researchers increased, the standard terminology to describe the census-defined geographic unit became “census tracts”.²

Health concerns nevertheless remained key to increasing the number of cities with census tracts. Of particular note were the efforts of Howard Whipple Green, a Cleveland health researcher keen on

creating small-area Public Health Nursing Districts with permanent boundaries (Figure 1).^{2, pp. 11-16}. With Laidlaw's advice, he was able to establish census tracts in Cleveland in 1927 and encouraged similar projects in other cities, such that the number of cities with census tracts jumped to 18 in 1930 census. In all cities, boundaries were established by local census tract committees, in keeping with census-defined standards, often with advice from the local public health agencies. Whipple also published several ground-breaking works on infant mortality rates and tuberculosis rates in relation to census tract socioeconomic and racial/ethnic composition.³⁻⁶ His methodology of combining census tract and public health data was subsequently adopted by other researchers to study socioeconomic and racial/ethnic disparities in mortality plus the incidence of tuberculosis and cancer; influential analyses include Nathan's 1932 investigation of "Health Conditions in North Harlem 1923-1927"⁷, Sheps and Watkins' 1947 study of "Mortality in the Socio-Economic Districts of New Haven"⁸, Terris' 1948 paper on "Relation of Economic Status to Tuberculosis Mortality by Age and Sex"⁹, and Cohart's 1954 article on "Socioeconomic Distribution of Stomach Cancer in New Haven."¹⁰ Together, this research opened the door to more widespread use of census tract data in the social sciences.^{2,11}

By the 1950s, census tracts were employed in research on health, fertility, immigration, marriage, social stratification, crime, employment, education, residential segregation, and urban ecology, as summarized in the American Statistical Association's monograph on the *Golden Anniversary of Census Tracts, 1956*² and the 1955 landmark publication of *Social Area Analysis* by Eshref Shevsky and Wendell Bell.¹¹ A major 1959 review paper on "Analysis of Vital Statistics by Census Tract," written by Elizabeth J. Coulter and Lillian Guralnick, published in the *Journal of the American Statistical Association*,¹² reported on results of a survey of official health agencies "on the current uses of census tract data with special references to studies of health as related to socioeconomic status." Results indicated not only strong interest and growing use of these data but also a desire to have census tracts delimited throughout the United States. Looking ahead, the article proposed that census tract data be used to develop a consistent socioeconomic measure that could be used nationwide, in order to "permit assessment of mortality rates between and within cities and serve to extend considerably an understanding of conditions

producing differential mortality” as well as “help in evaluation of other factors in mortality such as air pollution, climate, or the impact of city services.”¹², pp. 738-739

The census tract in the 21st century: serving the body politic – and revealing health inequities

Jumping ahead to the start of the 21st century, several of the goals articulated by Coulter and Guralnick have been accomplished. In 2000, census tracts were for the first time defined for the entire US.¹³ Scientific use of census tract data continues to expand in the myriad disciplines constituting what may be termed the population sciences, embracing health, social, and geographic sciences alike; listed in alphabetical order, they include: anthropology, criminology, demography, ecology, economics, education, environmental sciences, epidemiology, geography, history, political science, public health, sociology, statistics, and urban planning, to name a few. Census tracts likewise widely serve as the fundamental unit employed by myriad public agencies to create local administrative areas and are used to designate diverse policy-related entities, including “Urban Empowerment Zones,” “Medically Underserved Areas,” “Qualified Census Tracts” for the low-income housing credit, and “poverty areas” (defined as regions where 20% or more of the population is below the US poverty line).¹³⁻¹⁶ Comprising a fundamental building block for public policy and public planning, in health and other sectors, census tracts have thus become a key component of the public data central to documenting need and allocating resources within the body politic.

As importantly, recent research has begun to accomplish Coulter and Guralnik’s hope for improving documentation and analysis of, and interventions to address, social inequalities in health and health care. New work, for example, has shown that the census tract poverty measure can validly be used for monitoring and analyzing US socioeconomic inequalities in health, thereby overcoming the problem posed by the lack of socioeconomic data in most US public health surveillance systems and hospital data systems.¹⁷⁻²⁵ Building on a research legacy extending back to Laidlaw and Whipple’s time, albeit with greatly improved computer technology,²⁶ the basic method is to characterize both cases and the population from which they arise by the socioeconomic characteristics of their immediate residential area,

i.e., the census tract. This is accomplished by geocoding the cases' residential address, obtained from their health records, to the corresponding census tract, whose population counts (e.g., by age, gender, and race/ethnicity) in turn are obtained from US census data. Rates of outcomes among persons living in, say, poorer census tracts can thus be compared to the corresponding rates among persons living in more affluent census tracts.

Importantly, new methodologic research on census tract socioeconomic measures has shown that they: (1) can be used with any population, regardless of age or employment status (and hence, unlike education and income, can be used with persons still in school or not earning an income, whether because they are unemployed or retired); (2) provide estimates of socioeconomic gradients on par with those obtained with individual-level socioeconomic measures (hence indicating ecologic fallacy is not, empirically, a major concern); and (3) compared to block group socioeconomic measures, have the advantage of yielding a higher proportion of addresses geocoded to the relevant census-defined geographic unit.¹⁷⁻²⁴ Additionally, they are not affected by problems unique to ZIP Code level data, e.g., the spatiotemporal mismatch of census and ZIP Code geographic boundaries, which has been shown to introduce sufficient bias to reverse, wrongly, the direction of the socioeconomic gradient detected using ZIP Code-level socioeconomic data.²⁵

Currently, work using geocoding and census tract area-based socioeconomic measures not only demonstrates the existence of widespread socioeconomic disparities in health within diverse US racial/ethnic groups for outcomes spanning from birth to death, but also underscores the contribution of socioeconomic inequity to racial/ethnic disparities in health.¹⁷⁻²⁸ Census tract data are likewise increasingly being used to analyze inequities in access to health care, providing insight into the contribution of economic and geographic barriers to obtaining adequate care.²⁸⁻³⁰ At a time when fully 70% of the 467 US public health objectives for *Healthy People 2010* lack socioeconomic targets because of a lack of socioeconomic data in the relevant health data sources, the methodology of geocoding and using census tract area-based socioeconomic measures provides a promising solution to the longstanding problem of unavailable routine data, at the national, state, and local level, on US socioeconomic health

and health care inequities.¹⁷⁻²⁴ Further underscoring the attractiveness of this approach, starting in 2010, the American Community Survey is scheduled to be releasing census tract data annually, based on five-year rolling estimates, thereby no longer necessitating reliance on only decennial data.³¹

Additionally, rising use of census tract data in the population sciences has also spurred growing awareness of – and empirical evidence documenting – the significance of social context in shaping the societal distribution of myriad outcomes, including population health.³²⁻³⁴ Specifically, multilevel analyses, incorporating both individual-level and census tract socioeconomic data, have shown that both levels independently, and in some cases synergistically, are associated with risk of diverse health outcomes, once again ranging from birth to death.^{21,32-35} The implications are that determinants of health cannot be analyzed exclusively at the individual level; context matters. As for what aspects of this context are germane: although much remains to be elucidated, studies indicate relevant aspects include, at one level, the presence of harmful or beneficial neighborhood physical characteristics, themselves in part reflecting zoning ordinances (e.g., liquor stores and fast-food restaurants versus supermarkets and parks; mixed-use neighborhoods versus neighborhoods without commercial or industrial sites, etc.), and, at another, the impact of concentrated poverty on risk of adverse social interactions, including violence.³²⁻³⁵ Tellingly, a recently released federal report, *Shaping a Vision of Health Statistics for the 21st Century*,³⁶ includes as its conceptual model a figure that literally places population health—and its distribution—at the center, with determinants framed in terms of “place & time,” “context,” and “community attributes.” In this model, population health is clearly shaped by the social and physical conditions of neighborhoods, workplaces, and society at large – and the use of census tract data has played an important part in revealing these connections. For this reason, the report recommended that “all ongoing health statistics data sets should be geocoded to at least the census block group level” – which is a subdivision of the census tract – thereby enabling linkage of the health data to census-derived community characteristics, including at the census tract level.

Centenary reflections: census tracts, governance, and social responsibility

In summary, the census tract centenary is grounds for both celebration and reflection. Celebration, because the census tract -- as a unit of geography at once spatial, social, and political -- has usefully thrived for a century. And reflection, because the census tract can achieve its objective precisely because it is a national creation of democratic governance informed by local input, in accord with uniform standards, rather than driven by the imperatives of short-term political gain (per the prior use of wards to tally populations) or market distortions. Laidlaw's envisioned areas, critical for his hoped for "convenient and scientific city map system"¹ -- have met and exceeded initial expectations. Census tracts not only provide a stable geographic unit for estimating the number and characteristics of the people and housing units located within them, but they have also expanded scientific understanding of the impact of context on the social patterning of the public's well-being, with obvious policy relevance. Census tracts thus stand as an important testament to the useful role of government -- and scientists -- in assuring the existence of population data necessary for understanding and improving the health and well-being of our bodies and the body politic. We do well to honor their 100th anniversary.

REFERENCES:

1. Laidlaw W. Federation districts and a suggestion for a convenient and scientific city map system. *Federation* 1906;IV:1-6.
2. Watkins RJ, Swift Jr A, Green HW, Eckler AR. *Golden Anniversary of Census Tracts, 1956*. Washington, DC: American Statistical Association, 1956.
3. Green HW. The use of census tracts in analyzing the population of a metropolitan community. *J Am Stat Assoc* 1933;28:147-153.
4. Green HW. *Tuberculosis and economic strata. Cleveland's Five-City Area, 1928-1931*. Cleveland, OH: Anti-Tuberculosis League, 1932.
5. Green HW. The use of census tracts in analyzing the population of a metropolitan community. *J Am Stat Assoc* 1933;28:147-153.
6. Green HW. *Infant mortality and economic status, Cleveland Five-City Area*. Cleveland, OH: Cleveland Health Council, 1939.
7. Nathan WB. *Health Conditions in North Harlem 1923-1927*. New York: National Tuberculosis Association, 1932.
8. Sheps C, Watkins JH. Mortality in the socio-economic districts of New Haven. *Yale J Biol Med* 1947-48;20:51-80.
9. Terris M. Relation of economic status to tuberculosis mortality by age and sex. *Am J Public Health* 1948;38:1061-1070.
10. Cohart EM. Socioeconomic distribution of stomach cancer in New Haven. *Cancer* 1954;8:455-461.
11. Shevky E, Bell W. *Social Area Analysis: Theory, Illustrative Application, and Computational Procedures*. Stanford, CA: Stanford University Press, 1955.
12. Coulter EJ, Guralnick L. Analysis of vital statistics by census tract. *J Am Stat Assoc* 1959;54:730-740.

13. US Census Bureau. *Census Bureau Geography*. Available at: <http://www.census.gov/geo/www/index.html> . Accessed on September 19, 2005.
14. Department of Housing and Urban Development. Qualified Census Tracts and Difficult Development Areas. Available at: <http://www.huduser.org/datasets/qct/qct99home.html>. Accessed September 19, 2005.
15. Health Resources and Services Administration. Health Professional Shortage Areas. Available at: <http://bhpr.hrsa.gov/shortage/>. Accessed September 19, 2005.
16. US Bureau of the Census. Poverty areas. Available at: <http://www.census.gov/population/socdemo/statbriefs/povarea.html>. Accessed September 19, 2005.
17. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: the *Public Health Disparities Geocoding Project*. *Am J Public Health* 2005;95:312-323.
18. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health* 1992;82:703-710.
19. Krieger N, Chen JT, Ebel G. Can we monitor socioeconomic inequalities in health? A survey of U.S. Health Departments' data collection and reporting practices. *Public Health Rep* 1997;112:481-91.
20. Rehkopf DH, Haughton L, Chen JT, Waterman PD, Subramanian SV, Krieger N. Monitoring socioeconomic disparities in death: comparing individual-level education and area-based socioeconomic measures. *Am J Public Health* (in press).
21. Subramanian S V, Chen JT, Rehkopf DR, Waterman PD, Krieger N. Comparing individual and area-based socioeconomic measures for the surveillance of health disparities: a multilevel analysis of Massachusetts (US) births, 1988-92. *Am J Epidemiol* (in press).
22. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice

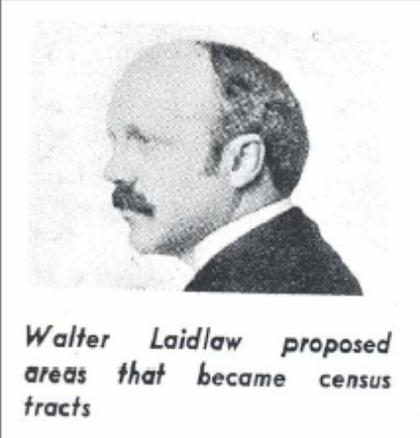
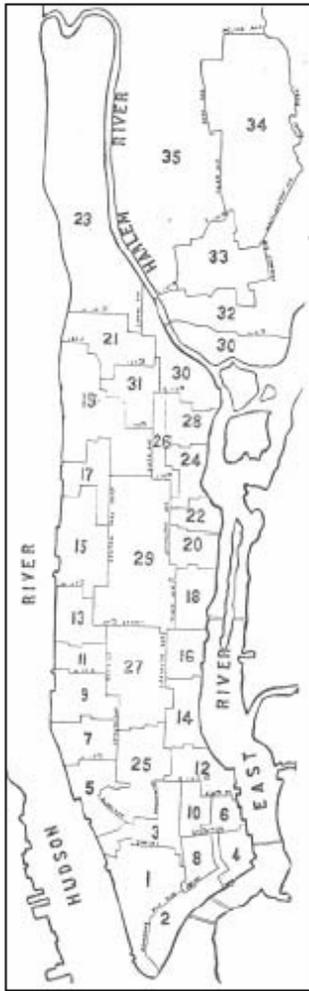
- of area-based measure and geographic level matter?: *The Public Health Disparities Geocoding Project*. *Am J Epidemiol* 2002;156:471-482.
23. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: *The Public Health Disparities Geocoding Project (US)*. *J Epidemiol Community Health* 2003;57:186-199.
 24. Krieger N, Waterman PD, Chen JT, Soobader MJ, Subramanian S. Monitoring Socioeconomic Inequalities in Sexually Transmitted Infections, Tuberculosis, and Violence: Geocoding and Choice of Area-Based Socioeconomic Measures -- *The Public Health Disparities Geocoding Project (US)*. *Public Health Rep* 2003;118:240-260.
 25. Krieger N, Waterman P, Chen JT, Soobader MJ, Subramanian SV, Carson R. Zip code caveat: bias due to spatiotemporal mismatches between zip codes and US census-defined geographic areas--*The Public Health Disparities Geocoding Project*. *Am J Public Health* 2002;92:1100-1102.
 26. Ricketts TC. Geographic information systems and public health. *Annu Rev Public Health* 2003; 24:1-6.
 27. Singh GK, Miller BA, Hankey BF, Edwards BK. *Area Socioeconomic Variations in U.S. Cancer Incidence, Mortality, Stage, Treatment, and Survival, 1975-1999*. NCI Cancer Surveillance Monograph Series, Number 4. Bethesda, MD: National Cancer Institute, 2003. NIH Pub. No. 03-5417.
 28. Braveman PA, Egerter SA, Cubbin C, Marchi KS. An approach to studying social disparities in health and health care. *Am J Public Health* 2004;94:2139-2148.
 29. Brooks-Gunn J, McCormick MC, Klebanov PK, McCarton C. Health care use of 3-year-old low birth weight premature children: Effects of family and neighborhood poverty. *J Pediatrics* 1998;132:971-975.

30. Phillips RL Jr, Kinman EL, Schnitzer PG, Lindbloom EJ, Ewigman B. Using geographic information systems to understand health care access. *Arch Family Med* 2000;9:971-978.
31. US Census Bureau. American Community Survey. Data release dates. Available at: <http://www.census.gov/acs/www/SBasics/DataDiss/RelDates.htm>. Accessed September 19, 2005.
32. O'Campo P. Invited commentary: advancing theory and methods for multilevel models of residential neighborhoods and health. *Am J Epidemiol* 2003;157:9-13.
33. Subramanian SV, Jones K, Duncan C. Multilevel methods for public health research. In: Kawachi I, Berkman L (eds). *Neighborhoods and Health*. Oxford: Oxford University Press 2003:65-111.
34. Diez-Roux AV. Multilevel analysis in public health research. *Annu Rev Public Health* 2000; 21:171-192.
35. Maantay J. Zoning, equity, and public health. *Am J Public Health* 2001;91:1033-1041.
36. Friedman DJ, Hunter EL, Parrish RG. *Shaping a Vision of Health Statistics for the 21st Century*. Washington, DC: Department of Health and Human Services Data Council, Centers for Disease Control and Prevention, National Center for Health Statistics, and National Committee on Vital and Health Statistics, 2002. Available at: <http://www.ncvhs.hhs.gov/hsvision/>. Accessed September 19, 2005.

FIGURE LEGEND

100 years of US census tracts: 1906-2006.

100 years of US census tracts: 1906-2006



New York City's Year 2000 census tracts



Proposed 1906 New York City "sanitary areas"

1910 census: 8 cities with "census districts"

1930 census: 18 cities with "census tracts"

2000 census: census tracts US-wide

