City Living, Healthy Living?

The relationship between density and Health in U.S. cities and counties

Ella N. B. Hartenian (2011), honors thesis in Environmental Biology & Sustainable Development
Robert L. Dorit, Department of Biological Sciences, Nicholas J. Horton, Department of Mathematics & Statistics

Smith College

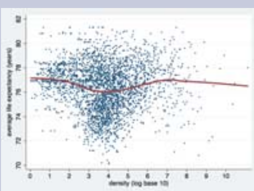
Introduction

Distributions of health outcomes take many different forms. The shape of the distribution depends on the particular health outcome and the specific population amongst other factors. In thinking about how to improve the health of populations it is particularly important to understand what influences health and specifically whose health it influences. Health outcomes also have an important geographical component; populations in certain areas often share certain risk factors and disease distributions. As urbanization continues and expands in the coming decades, understanding how place of residence impacts health will become increasingly important.

Here I look at several specific health outcomes and their relationship with socioeconomic status (SES) and density, which represents a rough measure of urbanicity.

Methodology and Data

In order to quantitatively analyze the impact of urbanicity on health, I selected several specific health outcomes (CHD & breast cancer) and looked at their distribution over a range of densities using county and city data. I then used multiple regression to test the relationship between the specific health outcomes, density and a range of measures of socioeconomic status.



While health is a composite variable, an overall measure of health like average life expectancy encompasses a wide range of factors. Choosing specific health outcomes allowed me to parse out the risk factors associated with each outcome and further investigate the nonlinear relationship seen here.

City-level data was obtained from the Big Cities Health Inventory

- Includes data on the 54 U.S. cities > 350,000 population
- Health outcomes in the report include: low birth weight, breast cancer mortality, coronary heart disease mortality, diabetes mortality

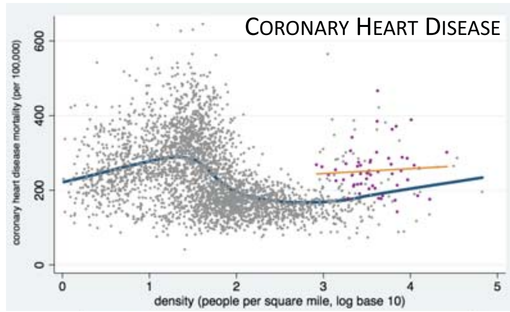
County-level data was obtained from the Community Health Status Indicators, the Department of Health and Human Services, the American FactFinder and County Health Rankings

- Includes data for all 3103 U.S. Counties
- And a range of health outcomes, measures of SES, health care prevision and access to care

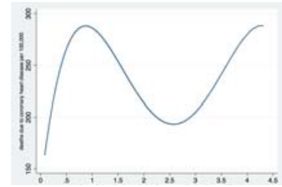
Thanks & Acknowledgement

First and foremost, thanks to Nick and Rob for all of their help and guidance relative to the project. I would also like to thank NACCHO for their release of the Big City Health Inventory dataset and the Tomlinson Fund for their support of this research.

Results and Analyses

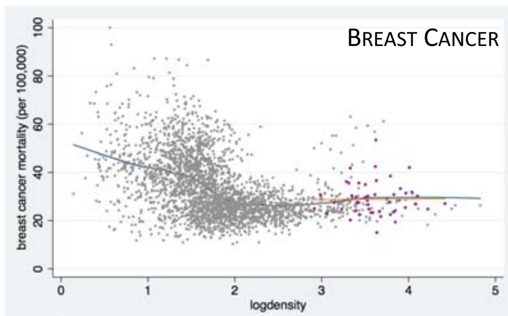


A scatter plot of logdensity and coronary heart disease mortality divided by city (purple) and county (grey) data with a blue locally weighted smoother of the county data and a yellow line of best fit for the city data.

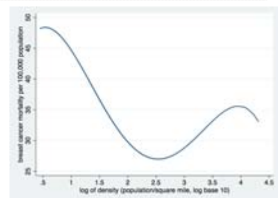


The predicted relationship between coronary heart disease mortality and density, controlling for a range of socioeconomic factors (see predictor/confounder table) when all variables are set to their mean value except density.

The regression model indicates a highly significant nonlinear poly-nomial that describes the relationship between CHD mortality, density and a range of socioeconomic factors ($F(11, 3087) = 188.98$; RMSE = 73.12; $p < 0.0001$). Compared to a similar model without density, log density explains an additional 7% of the variance ($R^2 = 0.40$ compared to an R^2 of 0.33).



A scatter plot of the relationship of log density with breast cancer mortality divided by city (purple) and county (grey) data with a blue locally weighted smoother for county data and a line of best fit in yellow for city data.



The predicted relationship between breast cancer mortality and density, controlling for a range of socioeconomic factors (see predictor/confounder table) when all variables are set to their mean value except density.

The regression model indicates a significant nonlinear polynomial that describes the relationship between breast cancer mortality and logdensity ($F(11,2706) = 110.40$; RMSE = 10.55; $p < 0.0001$). In this model, density explains about 13% of additional variance ($R^2 = 0.31$ compared to an R^2 of 0.18) relative to a model without log density as a predictor variable.

...tors and Confounders ...counties & county ...lents = 3,103)	Mean	Minimum	1st quartile	Median	3rd quartile	Maximum
... (population / sq mile)	240	0.04	17.56	43.21	105.29	66,940.07
...nsity)	1.67	0.02	1.27	1.65	2.03	4.83
...th weight (per 1,000 live births)	7.8	1.80	6.50	7.60	8.90	15.70
... of the 25+ population with at ...high school diploma	77	34.70	71.20	79.10	83.90	97.00
...efficient of inequality (0-100 ...100=greatest inequality)	43	32.90	41.00	43.00	45.70	60.00
... of families with incomes below ...erty line	11	1.00	6.70	9.50	13.30	55.70
... of the population over 65	16	2.90	12.70	15.10	17.80	36.20
... family income (rounded to ... thousand)	$42,000	14,000.00	36,000.00	41,000.00	47,000.00	97,000.00
... care physician rate (per ...)	57.33	0.00	30.10	50.90	75.20	621.90
...nes						
...ry Heart Disease Mortality ...99)	236.24	40.20	162.30	216.30	299.30	644.90
...Cancer Mortality (n=2,718)	33.21	10.10	23.80	29.50	41.30	99.90
...s (n=2,712)	7.81	0.50	5.90	7.50	9.45	20.80

log(density)	density (population/mile²)	characteristic counties (actual density)
0	0	Yukon-Koyukuk, AL (0.04) Northwest Arctic, AL (0.2)
0.5	2	Briscoe, TX (1.99) Kidder, ND (2.04) Crook, WY (2.06)
1	10	Wasco, OR (9.99) Kiowa, OK (10.08) Chaves, NM (10.11)
1.5	30	Fremont, CO (30.1) Fayette, IA (30.11) Clay, WV (30.17)
2	100	Tioga, NY (99.84) Angelina, TX (99.97) Christian, KY (100.19)
2.5	315	Buncombe, NC (315) Cumberland, ME (318) Mobile, AL (324)
3	1,000	Morris, NJ (1,002) Duval, FL (1,006) Wyandotte, KS (1,042)
3.5	3,160	Cuyahoga, OH (3,040) Ramsey, MN (3,281) Pinellas, FL (3,292)
4	10,000	Alexandria City, VA (8,452) Philadelphia, PA (11,233) Suffolk, MA (11,788)
4.5	31,620	Queens (20,409) Bronx, NY (31,709) Kings, NY (34,916)
5	100,000	New York, New York (66,940)

Discussion

- The relationships between coronary heart disease (CHD) mortality and breast cancer mortality and density is distinctly nonlinear.
- In the mid-low ranges of density there is a precipitous drop in both breast cancer and CHD mortality followed by a subsequent increase/ leveling off.
- This needs further attention and explanation; what other variables are influencing this relationship given that many aspects of SES are controlled for?
- City data are situated at the dense end of the county distribution and resemble the distributions of both health outcomes at the county level.
- Even though county and city data does not describe the same populations, these cities seem to be experiencing a similar overall phenomenon; after a density of 1,000 people/mile², density has little to no effect on CHD and breast cancer mortality.

Fifty four U.S. cities with populations greater than 350,000

U.S. counties organized by their 2003 urban influence code