# Are Distance Based on Addresses Better Than ZIP Codes for Assessing Geographic Access to Cancer Treatment?

Tzy-Mey (May) Kuo, PhD MPH

Integrated Cancer Information and Surveillance System

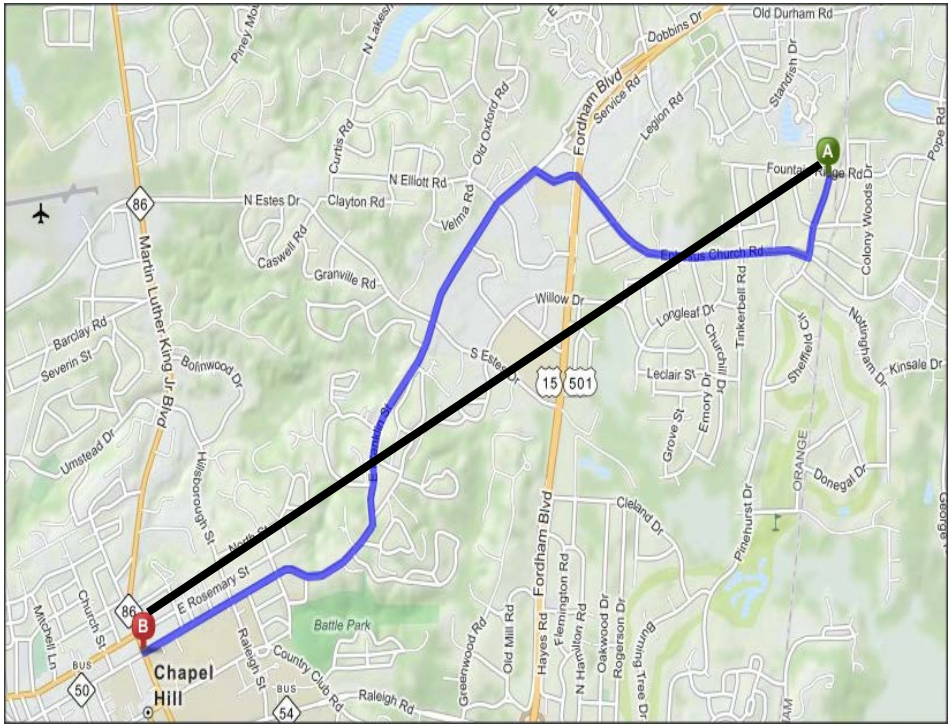
UNC LINEBERGER

ICISS

# Overview

- What distances are commonly used in public health studies?
- Geocoding
- Geographic coordinates
- Why do we care about distance?
- A study comparing distance effect on cancer care

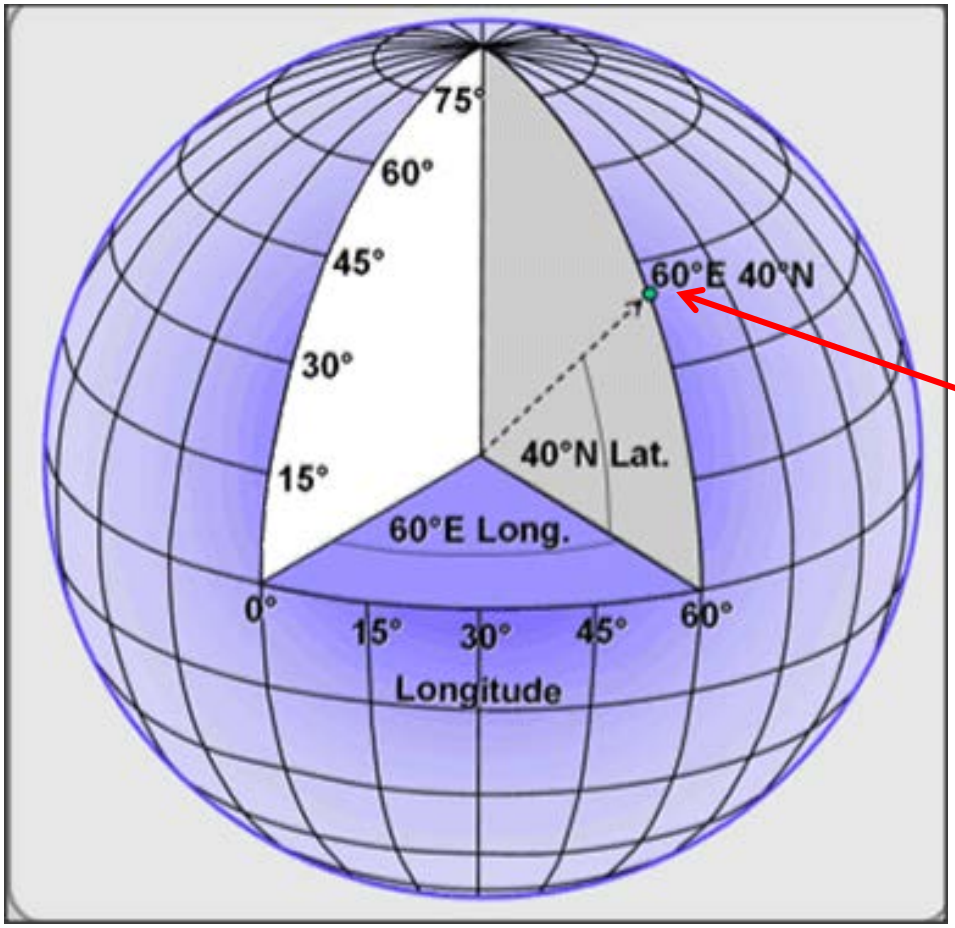
# What Distances Are Commonly Used?




Driving distance, network distance

Straight-line distance, Euclidean distance

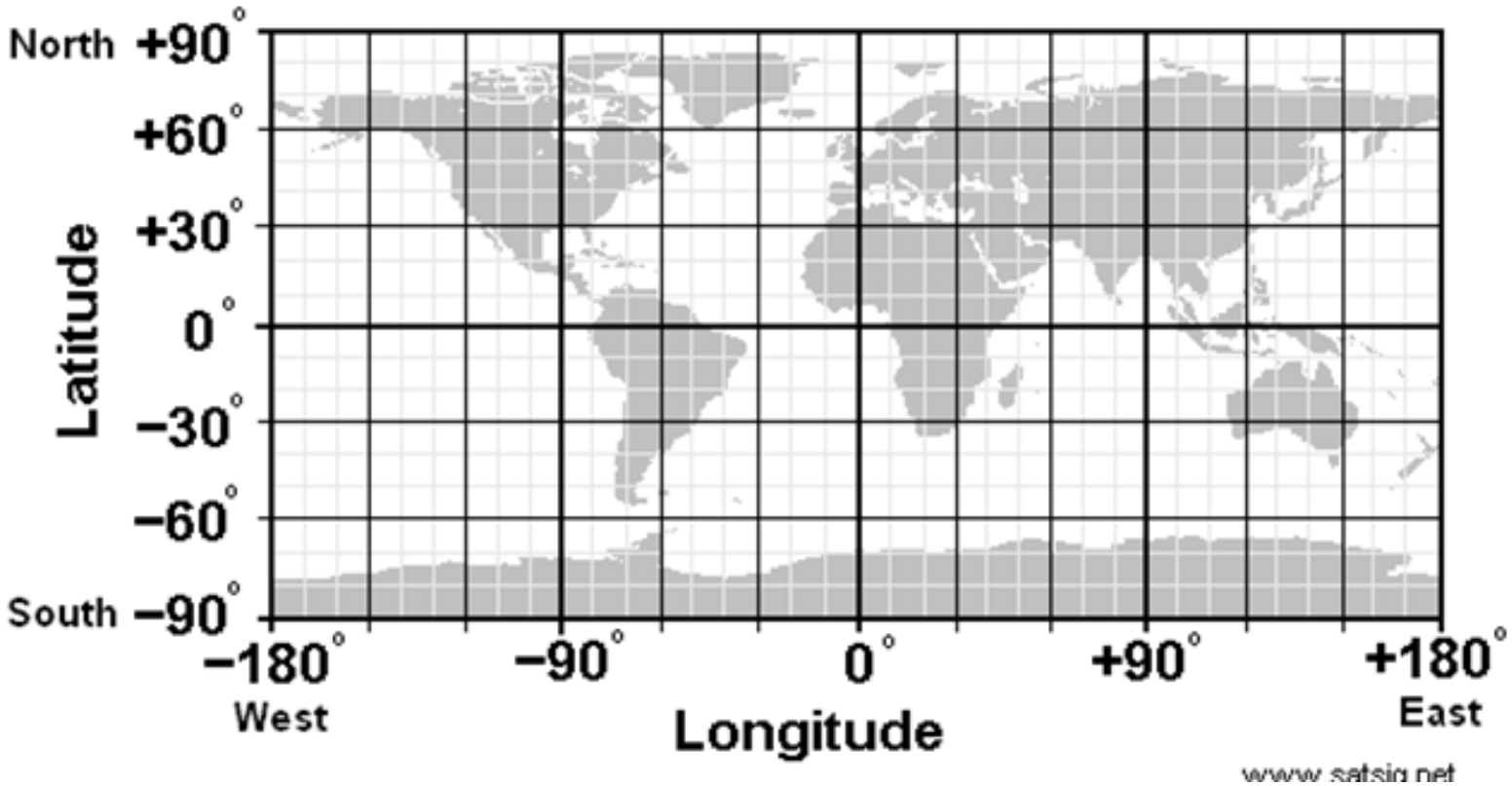
# Geocoding




Assign latitude and longitude for a location on the Earth's surface

Lat: 40 $^{o}$ North
Long: 60$^{o}$ East

# Geographic Coordinates: (Latitude, Longitude)

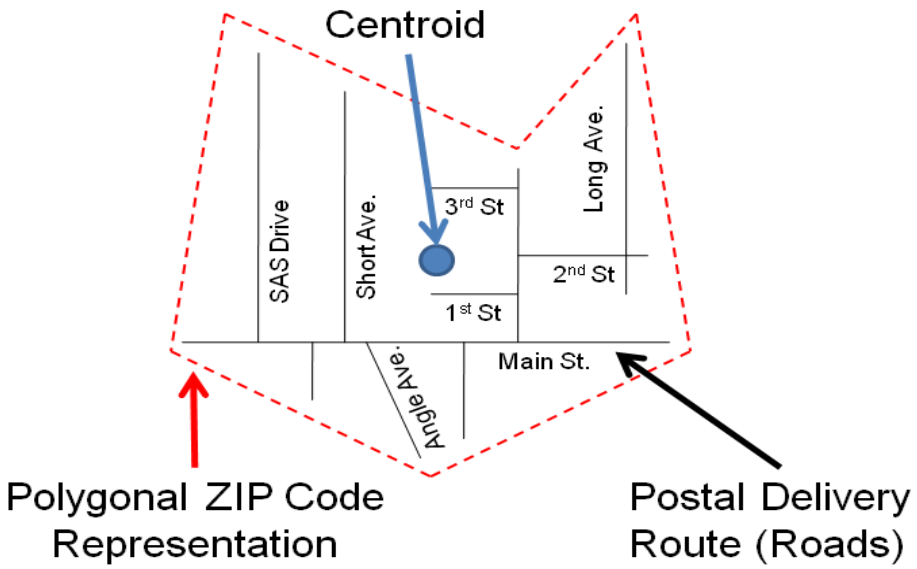


# Geocoding: Examples

- 101 E. Weaver Street, Carrboro, NC 27510

  ➡ (35.9110, -79.0720)

- ZIP code?




Massengill et al, Paper in SAS Global Forum 2010

# Why Do We Care About Distance?

- Distance as a proxy measure for geographic access to healthcare
- Findings about distance effect are inconsistent
  - Distance measures vary across studies
  - Geography (urban-rural) complicates the findings

# Why Do Urban and Rural Areas Matter?


# Study Objectives

- Compare distance effect on health care utilization:
  - Straight-line and driving distance
  - Addresses and ZIP codes data
- Examine distance effects across urban-rural areas

# Study Design

- Retrospective cohort:
  - Women age 65+, diagnosed with breast cancer in 2003-2005
  - Had breast conserving surgery
  - Needed to follow with radiation therapy (RT), usually given 5 days a week for about 5-6 weeks

# Data Sources

- NC Central Cancer Registry linked to Medicare claims
  - Cancer patients data and geocode addresses
  - All physicians who provided radiation treatment
- Medicare Physician Identification & Eligibility Records (MPIER)
  - Linked to physicians' data for providers' addresses
- Rural-Urban Commuting Area (RUCA) Data
  - Rural and urban indicator at ZIP code
- Area Resource File
  - Area-level control variables

# Distance Measures: Computation

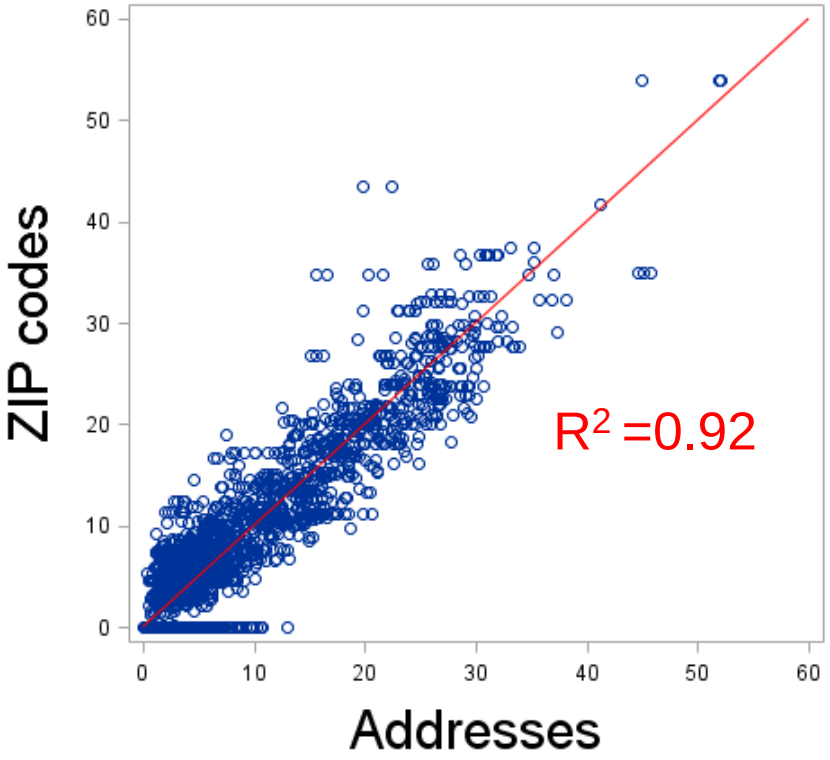
- Geocoded addresses and ZIP codes
- Nearest distance between patient's residence to radiation treatment physician
  - Straight-line distance: SAS GEODIST function
  - Driving distance: ESRI ArcGIS 10.1
    - Straight-line distance, addresses
    - Straight-line distance, ZIP code centroids
    - Driving distance, addresses
    - Driving distance, ZIP code centroids

# Statistical Analysis

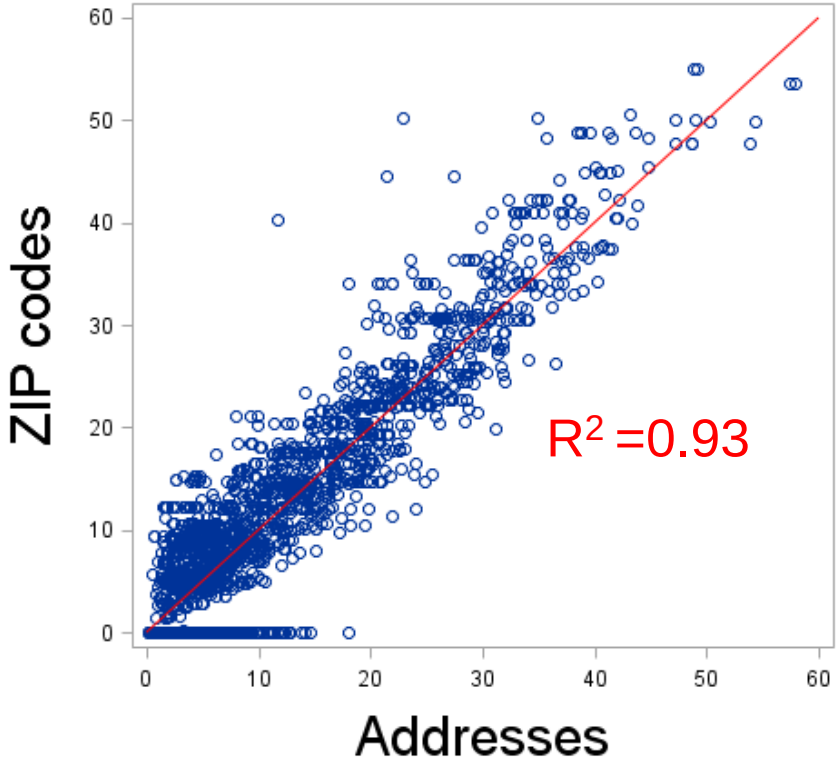
- Logistic Regression Models
  - Outcome variable: receipt of RT (yes vs. no)
  - Key independent variables: distance (logged), urban indicator, and their interactions
  - Control variables: patient demographics, comorbidities, tumor characteristics, county-level racial composition, median household income, and population density
  - Generalized Estimating Equation for clustered data

# Results: Correlations between Addresses and ZIP Codes

Straight-line Distance




Driving Distance



# Table 1. Descriptive Data among Different Distance Measures (N=1,938)

| | Straight-line Distance (Miles) | | Driving Distance (Miles) | |
|---|---|---|---|---|
| | Address | ZIP Centroid | Address | ZIP Centroid |
| Mean | 9.96 | 9.97 | 12.3 | 12.9* |
| Median | 6.55 | 7.65 | 8.27 | 9.93 |
| Min, Max | 0.1, 52.0 | 0, 53.9 | 0.1, 72.7 | 0, 74.6 |

* $p < 0.05$

- Mean straight-line distances were similar between ZIP codes and addresses
- Mean driving distance from ZIP codes was longer than that from addresses

# Table 2. Coefficient Estimates from Regression Models[a]

| | Straight-line Distance | | Driving Distance | |
|---|---|---|---|---|
| | **Address** | **ZIP code** | **Address** | **ZIP code** |
| **Urban vs. Rural** | 0.27 | 0.76 | 0.10 | 0.72 |
| **Linear Distance** | -0.24 | 0.02 | -0.34 | 0.01 |
| **Quadratic Distance** | 0.07 | 0.01 | 0.09 | 0.01 |
| **Linear Distance x Urban** | 0.38 | -0.16 | 0.51 | -0.14 |
| **Quadratic Distance x Urban** | -0.16* | -0.02 | -0.17* | -0.02 |

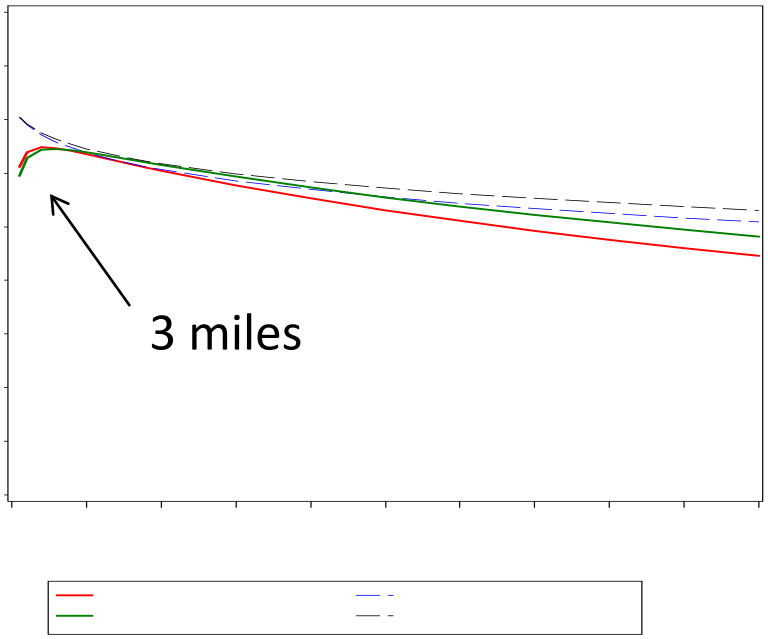
a:Models also included demographics, tumor characteristics, county level racial composition, median household income, and population density.

* $p<0.05$

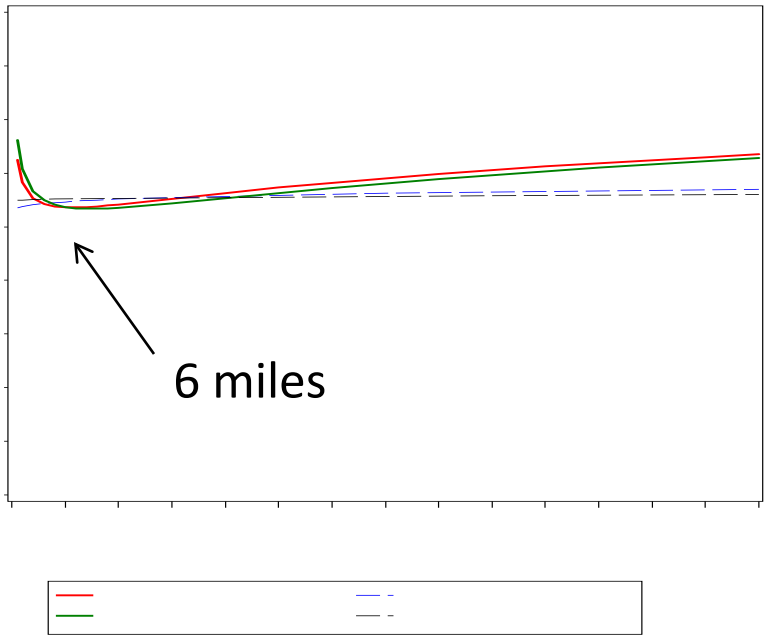
- Estimates for the key independent variables were similar between models using ZIP codes
- Estimates for the key independent variables using address showed a significant interaction effect

# Predicted Probabilities by Distance and Urban-Rural Areas

Urban Areas




Rural Areas




- **Probability of receiving RT increased by distance using addresses and then decreased (after 3 miles)**
- **Probability of receiving RT decreased by distance using addresses and then increased (after 6 miles)**

# Summary of Findings

- Same granularity of data (ZIP code and address) produced similar results
- Distance effect differed by urban-rural areas, found in address data only
- Findings about person level factors were consistent with literature: older, not married, low-income, more comorbidities, and earlier stage cancer diagnosis were associated with lower probability of receiving RT

# Conclusion

- Are distance based on addresses better than ZIP codes for assessing geographic access to cancer treatment?
- YES!

# Limitations and Future Studies

- Results may not be generalizable to:
  - Other states
  - Women <65 years old
- Limitations of claims data
- Further examination of women in rural area
  - Stayed in nearby areas to receive RT

# Acknowledgements


- Anne-Marie Meyer, Stephanie B. Wheeler, Brian Frizzelle, Marc Peterson, Adrian Meyer

- Work on this study was supported by the Integrated Cancer Information and Surveillance System (ICISS), UNC Lineberger Comprehensive Cancer Center with funding provided by the University Cancer Research Fund (UCRF) via the State of North Carolina.

# Thank you!

Questions?

tkuo@email.unc.edu